

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 395 842**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12P 21/00 (2006.01)

C07K 14/47 (2006.01)

C07K 16/00 (2006.01)

A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.01.2007 E 07718188 (1)**

97 Fecha y número de publicación de la concesión europea: **03.10.2012 EP 1974040**

54 Título: **Composiciones y métodos para la humanización y optimización de N-glicanos en plantas**

30 Prioridad:

17.01.2006 US 759298 P

07.04.2006 US 790373 P

11.04.2006 US 791178 P

09.06.2006 US 812702 P

11.08.2006 US 836998 P

21.11.2006 US 860358 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.02.2013

73 Titular/es:

SYNTHON BIOPHARMACEUTICALS B.V. (100.0%)

Microweg 22 PO Box 7071

6503 GN Nijmegen, NL

72 Inventor/es:

DICKEY, LYNN, F.;

COX, KEVIN, M. y

PEELE, CHARLES, G.

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 395 842 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Composiciones y métodos para la humanización y optimización de N-glicanos en plantas

Campo de la invención

La presente invención se refiere al campo de la biología molecular de plantas, más particularmente al campo de la producción de proteínas de mamífero recombinantes en plantas.

Antecedentes de la invención

Se han seleccionado como diana una serie de especies vegetales para su uso en el "cultivo molecular" de proteínas de mamífero de interés farmacéutico. Estos sistemas de expresión en plantas proporcionan una producción de bajo coste de proteínas de mamífero biológicamente activas y son fácilmente susceptibles de un escalado rápido y económico (Ma et al. (2003) Nat. Rev. Genet. 4: 794-805; Raskin et al. (2002) Trends Biotechnol. 20: 522-531). Un número grande de proteínas de mamífero y de plantas requieren un procesamiento post-traducciona para un plegamiento, un ensamblaje y un funcionamiento adecuados. De estas modificaciones, las diferencias en los modelos de glicosilación entre plantas y mamíferos ofrecen un reto para la posibilidad de que los sistemas de expresión de plantas produzcan proteínas de mamífero recombinantes de alta calidad para uso farmacéutico.

Según los péptidos se desplazan a través del retículo endoplasmático (RE) y los compartimentos de Golgi subcelulares, las cadenas de residuos de azúcares, o glicanos, se unen, dando lugar finalmente a la formación de glicoproteínas. El enlace entre las cadenas de azúcares y el péptido se produce mediante la formación de un enlace químico con solo uno de los cuatro aminoácidos proteicos: asparagina, serina, treonina e hidroxilisina. En base a este modelo de enlace, se han reconocido dos tipos básicos de cadenas de residuos de azúcares en las glicoproteínas: la cadena de azúcar ligada mediante *N*-glicósido (también denominada glicano *N*-ligado o *N*-glicano), que se une a los residuos de asparagina del péptido; y la cadena de azúcar ligada mediante *O*-glicósido, que se une a los residuos de serina, treonina e hidroxilisina del péptido.

Las cadenas de azúcares ligadas mediante *N*-glicósido, o *N*-glicanos, presentan diversas estructuras (véase, por ejemplo, Takahashi, ed. (1989) Biochemical Experimentation Method 23 – Method for Studying Glycoprotein Sugar Chain (Gakujutsu Shuppan Center), pero comparten un núcleo oligomanosídico común (véase la Figura 29A). Las etapas iniciales de la ruta de glicosilación que conducen a la formación de *N*-glicanos se conservan en plantas y animales. Sin embargo, las etapas finales implicadas en la formación compleja de *N*-glicanos difieren (Lerouge et al. (1998) Plant Mo. Biol. 38: 31-48; Steinkellner y Strasser (2003) Ann. Plant Rev. 9: 181-192). Las plantas producen glicoproteínas con *N*-glicanos complejos que tienen un núcleo que porta dos residuos de *N*-acetilglucosamina (GlcNAc) que es similar al observado en mamíferos. Sin embargo, en las glicoproteínas vegetales dicho núcleo está sustituido por un residuo de xilosa β 1,2-ligada (xilosa de núcleo), residuo que no se da en humanos, epítomos de Lewis^a y una fucosa α 1,3-ligada (α [1,3]-fucosa de núcleo) en lugar de una fucosa α 1,6-ligada de núcleo como en los mamíferos (véase, por ejemplo, Lerouge et al. (1998) Plant Mol. Biol. 38: 31-48 para una revisión) (véase también la Figura 29B). Se ha publicado que tanto los residuos de α (1,3)-fucosa como los de β (1,2)-xilosa son, al menos en parte, responsables de la inmunogenicidad de las glicoproteínas vegetales en mamíferos (véase, por ejemplo, Ree et al. (2000) J. Biol. Chem. 15: 11451-11458; Bardor et al. (2003) Glycobiol. 13: 427-434; Garcia-Casado et al. (1996) Glycobiol. 6: 471-477). Por lo tanto, la eliminación de estos residuos de azúcar potencialmente alergénicos de las glicoproteínas de mamífero producidas recombinantemente en plantas solucionaría las preocupaciones sobre el uso de dichas proteínas como productos farmacéuticos para el tratamiento de humanos.

Actualmente una serie de glicoproteínas producidas recombinantemente actúan como agentes terapéuticos o están bajo investigación clínica. Los ejemplos incluyen los interferones (IFNs), la eritropoietina (EPO), el activador de plasminógeno de tejido (tPA), la antitrombina, el factor de estimulación de colonia de granulocito-macrófago (GM-CSF) y los anticuerpos monoclonales terapéuticos (mABs). El componente de oligosacárido de las estructuras de *N*-glicano de las glicoproteínas puede influir en su eficacia terapéutica, así como en su estabilidad física, su resistencia al ataque de proteasas, su farmacocinética, su interacción con el sistema inmune y su actividad biológica específica. Véase, por ejemplo, Jenkins et al. (1996) Nature Biotechnol. 14: 975-981. Strasser et al. (2003) FEBS LETTERS: 561: 132-136) describe plantas de *Arabidopsis* con un modelo de *N*-glicosilación alterado, que se caracteriza por una reducción de la unión de residuos de α 1,3-fucosa y β 1,2-xilosa a los *N*-glicanos de las glicoproteínas.

Se necesitan otros métodos adicionales para alterar el modelo de glicosilación en sistemas de expresión vegetales, específicamente para inhibir la glicosilación específica de planta de la estructura de núcleo eucariótico, para producir de forma ventajosa proteínas de mamífero recombinantes con un modelo de glicosilación humanizado.

Breve resumen de la invención

Se proporcionan métodos para alterar el modelo de *N*-glicosilación de proteínas en plantas superiores. Los métodos comprenden transformar de manera estable la planta con al menos una construcción de nucleótido recombinante que proporcione la inhibición de la expresión de α 1,3-fucosiltransferasa (FucT) y de β 1,2-xiloxiltransferasa (XylIT) en una planta. El uso de estas construcciones para inhibir o suprimir la expresión de uno o ambas enzimas, y sus isoformas, proporciona de forma ventajosa la producción de proteínas endógenas y heterólogas que tienen un

modelo de *N*-glicosilación "humanizado" sin afectar al crecimiento y al desarrollo de la planta. Se proporcionan las plantas superiores transformadas de manera estable que presentan este modelo de *N*-glicosilación de proteína. La planta es una lenteja de agua, es decir, un miembro de la familia *Lemnaceae*, por ejemplo, una *Lemna sp.*

Las plantas transgénicas de la invención tienen la capacidad de producir proteínas de mamífero recombinantes que tienen un modelo de *N*-glicosilación que es más similar al del hospedante mamífero. Por tanto, en algunas realizaciones, las proteínas de mamífero producidas recombinantemente son glicoproteínas que comprenden *N*-glicanos que presentan una reducción en la unión de los residuos de $\alpha(1,3)$ -fucosa y $\beta(1,2)$ -xilosa específicos de plantas. En otras realizaciones, estas glicoproteínas producidas recombinantemente comprenden *N*-glicanos complejos que están desprovistos de estos residuos específicos de plantas. En otras realizaciones adicionales, estas glicoproteínas producidas recombinantemente tienen $\text{GlcNAc}_2\text{Man}_3\text{GlcNAc}_2$ como única especie de glicano unida a el(los) sitio(s) de glicosilación de asparagina dentro de la glicoproteína.

En algunas realizaciones, la glicoproteína producida recombinantemente es un anticuerpo monoclonal. De este modo, la presente invención proporciona plantas transgénicas que son capaces de producir anticuerpos monoclonales de glicano optimizado que se unen específicamente a una proteína diana de interés, que incluyen anticuerpos monoclonales que tienen una función efectora aumentada. Por tanto, en algunas realizaciones, los anticuerpos monoclonales de glicano optimizado producidos recombinantemente comprenden *N*-glicanos complejos que presentan una reducción en la unión de residuos de fucosa $\alpha(1,3)$ -ligados, aumentando con ello la actividad ADCC de dichos anticuerpos. En otras realizaciones, los anticuerpos monoclonales producidos recombinantemente comprenden glicanos complejos *N*-ligados que están desprovistos de dichos residuos de fucosa específicos de plantas. De esta manera, la presente invención proporciona la producción de una composición de anticuerpos monoclonales, en donde al menos el 90% o más del anticuerpo intacto está representado por una única glicoforma, más particularmente, por la glicoforma G0. Por tanto, en algunas realizaciones de la invención, los anticuerpos monoclonales producidos recombinantemente presentan una función efectora incrementada, en donde la actividad ADCC aumenta y/o la relación de actividad ADCC/CDC aumenta. En algunas de estas realizaciones, los anticuerpos monoclonales producidos recombinantemente presentan una actividad CDC disminuida, lo que puede reducir de manera ventajosa el potencial de efectos secundarios adversos relacionados con la activación CDC tras la administración. Estos anticuerpos monoclonales de glicano optimizado se pueden usar de manera ventajosa para alterar las rutas actuales de administración y los regímenes terapéuticos actuales, ya que su función efectora aumentada significa que se pueden dosificar a concentraciones más bajas y con menor frecuencia, reduciéndose así el potencial de toxicidad del anticuerpo y/o de desarrollo de tolerancia al anticuerpo. Adicionalmente, su función efectora mejorada da lugar a nuevas estrategias de tratamiento de indicios clínicos que previamente se han mostrado resistentes o refractarios al tratamiento con el anticuerpo monoclonal correspondiente producido en otros sistemas hospedantes recombinantes.

También se describen las composiciones para llevar a la práctica los métodos de la invención. Las composiciones comprenden nuevos polinucleótidos y polipéptidos aislados que codifican $\alpha 1,3$ -fucosiltransferasa y $\beta 1,2$ -xilosiltransferasa de *Lemna minor*, y variantes y fragmentos de los mismos. Se describen las construcciones de nucleótidos recombinantes dirigidas a la expresión de estas dos proteínas, o a la expresión de sus variantes, al igual que se describen células vegetales, tejidos vegetales, plantas y semillas que comprenden dichas construcciones recombinantes.

Por lo tanto la invención proporciona una planta de lenteja de agua o una célula o nódulo de planta de lenteja de agua que expresa una glicoproteína recombinante que tiene un perfil de *N*-glicosilación sustancialmente homogéneo, en donde al menos el 90% de las especies de *N*-glicano presentes en el perfil son glicanos G0, y en donde dicha planta, célula o nódulo vegetal comprende al menos una construcción de nucleótido recombinante que proporciona la inhibición de la expresión de $\alpha 1,3$ -fucosiltransferasa (FucT) y $\beta 1,2$ -xilosiltransferasa (XylT) en dicha planta, célula o nódulo vegetal por medio de:

(a) supresión o co-supresión sentido a través de una casete de expresión que expresa una molécula de ARN correspondiente a todo o a una parte del ARN mensajero que codifica una FucT o una XylT, o ambas, en la orientación sentido;

(b) supresión antisentido a través de una casete de expresión que expresa una molécula de ARN complementaria a todo o a una parte del ARN mensajero que codifica una FucT o una XylT, o ambas, en la orientación antisentido;

(c) interferencia de ARN de cadena doble, en donde el ARN sentido de (a) y el ARN antisentido de (b) son expresados en la misma célula;

(d)

(i) interferencia de ARN de horquilla (hpARN), o

(ii) interferencia de ARN de horquilla que contiene intrón (ihpARN)

a través de una casete de expresión de ARNi; o

(e) inhibición de microARN (miARN) a través de una construcción que expresa un ARN que forma una estructura de horquilla que contiene una secuencia de 22 nucleótidos que es complementaria a una secuencia diana de FucT o de XylT y que contiene 22 nucleótidos de dicha secuencia de FucT o de XylT en orientación sentido y 21 nucleótidos de una

5 (f) secuencia antisentido correspondiente que es complementaria a la secuencia sentido.

La invención también proporciona un método para usar la planta de lenteja de agua o la célula o nódulo vegetal de planta de lenteja de agua de la invención para producir una composición de glicoproteínas que presenta un perfil de *N*-glicosilación sustancialmente homogéneo, en donde al menos un 90% de las especies de *N*-glicano presentes en el perfil son glicanos G0, comprendiendo dicho método:

10 (a) cultivar dicha planta de lenteja de agua o dicha célula o nódulo de planta de lenteja de agua en las condiciones adecuadas para la expresión de dicha glicoproteína; y,

(b) recolectar dicha glicoproteína de dicha planta de lenteja de agua, célula o nódulo vegetal, o dicho medio de cultivo;

produciendo con ello dicha composición de glicoproteína.

15 La glicoproteína preferiblemente es un anticuerpo monoclonal.

Breve descripción de las figuras formales

Figura 1: establece las secuencias de ADN (SEC ID N°: 1, secuencia codificadora establecida en la SEC ID N°: 2) y de aminoácidos (SEC ID N°: 3) correspondientes a la α 1,3-fucosiltransferasa (FucT) de *Lemna minor*. La secuencia codificadora se muestra en negrita. Los nucleótidos denotados por un subrayado sencillo (-) corresponden al fragmento directo de FucT dentro de la casete de expresión de ARNi diseñada para inhibir la expresión de FucT (véase la Figura 5); los nucleótidos denotados mediante el subrayado doble (=) corresponden a la secuencia espaciadora dentro de la casete de expresión de ARNi. El fragmento inverso de FucT de la casete de expresión de ARNi es el antisentido del fragmento directo de FucT mostrado aquí.

Figura 2: establece un alineamiento de la FucT de *Lemna minor* de SEC ID N°: 3 con α 1,3-fucosiltransferasas de otras plantas superiores.

Figura 3: establece la secuencia de ADN (SEC ID N°: 4; secuencia codificación establecida en la SEC ID N°: 5) correspondiente a la β 1,2-xilosiltransferasa (XylT) isoforma n°1 y la secuencia de aminoácido codificada (SEC ID N°: 6). Los nucleótidos denotados por un subrayado sencillo (-) corresponden al fragmento directo de XylT dentro de la casete de expresión de ARNi diseñada para inhibir la expresión de XylT (véase la Figura 6); los nucleótidos denotados mediante el subrayado doble (=) corresponden a la secuencia espaciadora dentro de la casete de expresión de ARNi. El fragmento inverso de XylT de la casete de expresión de ARNi es el antisentido del fragmento directo de XylT mostrado aquí.

Figura 4: establece un alineamiento de la XylT de *Lemna minor* de SEC ID N°: 6 con β 1,2-xilosiltransferasas de otras plantas superiores.

Figura 5: establece una estrategia para diseñar un bloqueo de ARNi de gen individual de FucT de *Lemna minor*.

Figura 6: establece una estrategia para diseñar un bloqueo de ARNi de gen individual de XylT de *Lemna minor* basado en la secuencia de ADN correspondiente a la isoforma n°1 de XylT.

Figura 7: establece una estrategia para diseñar un bloqueo de ARNi de gen doble de FucT y XylT de *Lemna minor* en donde la porción de XylT del bloqueo de ARNi se basa en la secuencia de ADN correspondiente a la isoforma n°1 de XylT.

Figura 8: muestra la construcción Fuc02 que comprende una casete de expresión de ARNi diseñada para el bloque de ARNi de gen individual de FucT de *Lemna minor*. La expresión de la secuencia inhibidora de FucT (denotada con las flechas de FucT directo y FucT inverso; véase la Figura 5) se lleva a cabo en un elemento de control de expresión ligado operativamente (denotado como AocsAocsAocsA-masPmas) que comprende tres secuencias de activación hacia arriba (Aocs) derivadas del gen de octopina sintasa de *Agrobacterium tumefaciens* ligado operativamente a un promotor derivado de un gen de manopina sintasa de *Agrobacterium tumefaciens* (AmasPmas). Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 9: muestra la construcción Xyl02 que comprende una casete de expresión de ARNi diseñada para el bloqueo de ARNi de gen individual de XylT de *Lemna minor*. La expresión de la secuencia inhibidora de XylT (denotada por las flechas de XylT directo y XylT inverso; véase la Figura 6) se lleva a cabo mediante el elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina

sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 10: muestra la construcción XF02 que comprende una casete de expresión de ARNi quimérico diseñada para el bloqueo de ARNi de gen doble FucT/XylT de *Lemna minor*. El ARN de horquilla se expresa como secuencia quimérica (un ARN de horquilla quimérico), en donde los fragmentos de los dos genes están fusionados y expresados como un transcrito. La expresión de la secuencia inhibidora de FucT/XylT (denotada por las flechas de FucT y XylT directa y de XylT y FucT inversa; véase la Figura 7) se lleva a cabo mediante el elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 11: muestra la construcción XF03 que comprende una casete de expresión de ARNi diseñada para el bloqueo de ARNi de gen doble de FucT/XylT de *Lemna minor*. La casete expresa dos horquillas de ARNi, una dirigida a la expresión de FucT, la otra dirigida a la expresión de XylT. La expresión de la secuencia inhibidora de FucT (denotada mediante flechas FucT directa y FucT inversa; véase la Figura 5) se lleva a cabo mediante un elemento de control de la expresión ligado operativamente que comprende el promotor de ubiquitina de *Lemna minor* más 5' UTR (promotor LmUbq) e intrón (intrón LmUbq) (véase la SEC ID N°: 7). La expresión de la secuencia inhibidora de XylT (denotada con flechas de XylT directa y XylT inversa; véase la Figura 6) se lleva a cabo a través del elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 12: muestra la construcción mAbI04 que proporciona la co-expresión de un anticuerpo monoclonal IgG 1 (referido en la presente memoria como mAbI) y el bloqueo de gen doble de FucT y XylT de *Lemna minor*, en donde se expresa un ARN de horquilla quimérico dirigido a la expresión de FucT y XylT. La expresión de la secuencia inhibidora de FucT/XylT (denotada mediante las flechas de FucT y XylT directa y de XylT y FucT inversa; véase la Figura 7) se lleva a cabo mediante un elemento de control de expresión ligado operativamente que comprende el promotor de ubiquitina de *Spirodella polyrrhiza* más 5' UTR (promotor SpUbq) e intrón (intrón SpUbq) (véase la SEC ID N°: 8). La expresión de la cadena ligera de IgG1 se lleva a cabo mediante el elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 13: muestra la construcción mAbI05 que proporciona la co-expresión de mAbI y el bloqueo doble de FucT y XylT de *Lemna minor*, en donde se expresan dos ARNs de horquilla, uno dirigido a la expresión de FucT, el otro dirigido a la expresión de XylT. La expresión de la secuencia inhibidora de FucT (denotada mediante flechas de FucT directa y FucT inversa; véase la Figura 5) se lleva a cabo mediante un elemento de control de expresión ligado operativamente que comprende el promotor de ubiquitina de *S. polyrrhiza* más 5' UTR (promotor SpUbq) e intrón (intrón SpUbq). La expresión de la secuencia inhibidora de XylT (denotada mediante flechas de XylT directa y XylT inversa; véase la Figura 6) se lleva a cabo mediante un elemento de control de expresión ligado operativamente que comprende el promotor de ubiquitina de *Lemna aequinoctialis* más 5' UTR (promotor LaUbq) e intrón (intrón LaUbq) (véase la SEC ID N°: 9). La expresión de la cadena ligera de IgG 1 se lleva a cabo mediante un elemento de control de expresión ligado operativamente que comprende el promotor de ubiquitina de *L. minor* más 5' UTR (promotor LmUbq) e intrón (intrón LmUbq). La expresión de la cadena pesada de IgG 1 se lleva a cabo mediante el elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*.

Figura 14: muestra la construcción de mAbI01 que proporciona la expresión de mAbI, en donde las expresiones de FucT y de XylT no son suprimidas. Las expresiones de la cadena ligera de IgG 1 y de la cadena pesada de IgG 1 se llevan a cabo independientemente mediante elemento de control de expresión ligado operativamente AocsAocsAocsAmasPmas. Líder RbcS, secuencia líder de subunidad pequeña rubisco; ADH1, intrón del gen de alcohol deshidrogenasa 1 de maíz; nos-ter, secuencia terminadora de nopalina sintasa (nos) de *Agrobacterium tumefaciens*. El mAbI01 es referido como construcción de mAbI01 "natural", ya que el mAbI expresado exhibe el perfil de glicosilación de la *L. minor* natural.

Figuras 15 y 16: muestran los datos del escrutinio primario de líneas vegetales de *L. minor* de ARNi transgénico que comprenden la construcción XF02 de la Figura 10.

Figura 17: muestra los datos del escrutinio primario de líneas vegetales de *L. minor* de ARNi transgénico que comprenden la construcción mAbI04 de la Figura 12 y la construcción mAbI05 de la Figura 13.

Figura 18: muestra la estructura y el peso molecular de *N*-glicanos mAb de *L. minor* natural derivatizados. "GnGn" representa la especie de *N*-glicano GlcNAc₂Man₃GlcNAc₂, también referida como especie de *N*-glicano G0. "GnGnX" representa la especie de *N*-glicano GlcNAc₂Man₃GlcNAc₂ con el residuo de β(1,2)-xilosa específico de planta unido. "GnGnXF" representa la especie de *N*-glicano GlcNAc₂Man₃GlcNAc₂ con los residuos específicos de planta β(1,2)-xilosa y α(1,3)-fucosa unidos.

Figura 19: muestra que la construcción de mAb101 natural (mostrada en la Figura 15) que proporciona la expresión del anticuerpo IgG1 monoclonal mAb1 en *L. minor*, sin supresión de ARNi de FucT y XylT de *L. minor*, produce un perfil de *N*-glicosilación con tres especies de *N*-glicano principales, que incluyen una especie que tiene la xilosa β 1,2-ligada y una especie que tiene ambos residuos, la xilosa β 1,2-ligada y la fucosa α 1,3-ligada de núcleo; este perfil se confirma con análisis de cromatografía de líquidos con espectrometría de masas (LC-MS) (Figura 20) y de MALDI (Figura 21).

Figura 22: muestra una superposición de las cantidades relativas de las diversas especies de *N*-glicano de mAb1 producidas en la línea de *L. minor* natural, que comprende la construcción mAb101 (sin supresión de FucT o XylT) y en las dos líneas transgénicas de *L. minor* que comprenden la construcción mAb104 de la Figura 12 (que proporcionan la co-expresión de mAb1 y la construcción de ARNi quimérico dirigidos a FucT y a X1T de *L. minor*). Cabe destacar el enriquecimiento de la especie de glicano GnGn (es decir, G0), sin residuos de xilosa β 1,2-ligada o de fucosa α 1,3-ligada de núcleo unidos, y la ausencia de especies que tengan residuos de xilosa β 1,2-ligada o de ambas, xilosa β 1,2-ligada y fucosa α 1,3-ligada de núcleo. "MGn" representa un precursor de *N*-glicano, en donde la estructura de núcleo de la trimanosa, Man₃GlcNAc₂, tiene una *N*-acetilglucosamina unida en el brazo de 1,3-manosa. "GnM" representa un precursor de *N*-glicano, en donde la estructura de núcleo de la trimanosa, Man₃GlcNAc₂, tiene una *N*-acetilglucosamina unida al brazo de 1,6-manosa. Dichos precursores de *N*-glicano representan cantidades traza del total de *N*-glicanos presentes en la muestra. Este perfil se confirma mediante análisis de espectrometría de masas (LC-MS) (Figura 23) y de MALDI (Figura 24).

Figura 25: muestra el análisis de masas intacto de composiciones de mAb1 producidas en *L. minor* natural (línea 20) que comprende la construcción mAb101. Cuando las expresiones de XylT y FucT no son suprimidas en *L. minor*, la composición de mAb1 producida recombinantemente es heterogénea, comprendiendo al menos 9 glicoformas diferentes, siendo la glicoforma G0XF³ la especie predominante de las presentes. Cabe destacar el pico muy pequeño que representa a la glicoforma G0.

Figura 26: muestra un análisis de masas intactas de composiciones de mAb1 producidas en *L. minor* transgénica (línea 15) que comprende la construcción mAb104 de la Figura 12. Cuando las expresiones de XylT y FucT son suprimidas en *L. minor* usando esta construcción de ARNi quimérica, la composición de mAb1 intacta es sustancialmente homogénea para *N*-glicanos G0, con solo cantidades traza de *N*-glicanos precursores presentes (representados por las especies de glicano precursor GnM y MGn). Adicionalmente, la composición de mAb1 es sustancialmente homogénea para la glicoforma G0, en donde ambos sitios de glicosilación están ocupados por especies de *N*-glicano G0, con tres picos menores que reflejan cantidades traza de glicoformas precursoras (un pico que muestra mAb1 que tiene una región Fc en donde el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de la otra cadena pesada está sin glicosilar; otro pico que muestra mAb1 que tiene una región Fc en donde el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de otra cadena pesada tiene el glicano precursor GnM o MGn unido a Asn 297, y otro pico que muestra mAb1 que tiene una región en donde el sitio de glicosilación de Asn 297 de cada uno de los dominios C_H2 tiene una especie de glicano unida, con una tercera especie de glicano G0 unida a un sitio de glicosilación adicional dentro de la estructura de mAb1).

Figura 27: muestra un análisis de masas intactas de las composiciones de mAb1 producidas en *L. minor* transgénica (línea 72) que comprende la construcción mAb105 de la Figura 13. Cuando las expresiones de XylT y de FucT son suprimidas en *L. minor* usando esta construcción, la composición de mAb1 intacta es sustancialmente homogénea para *N*-glicanos G0, solo con cantidades traza de especies de *N*-glicano presentes (representadas por las especies de glicano precursor GnM y MGn). Adicionalmente, la composición de mAb1 es sustancialmente homogénea (al menos un 90%) para la glicoforma G0, con los mismos tres picos menores que reflejan las glicoformas precursoras obtenidos para la construcción mAb104.

Figura 28: resume dos posibles diseños para dirigir la expresión de genes FucT y XylT individuales.

Figura 29A: muestra la estructura de núcleo oligomanosídico común de *N*-glicanos complejos de glicoproteínas producidas en plantas y animales. En mamíferos, la estructura de núcleo puede incluir un residuo de fucosa en el que la posición 1 de la fucosa está ligada a la posición 6 de la *N*-acetilglucosamina en el extremo reductor a través de un enlace α (es decir, fucosa α (1,6)-ligada).

Figura 29B: muestra las modificaciones específicas de planta de estos *N*-glicanos. Los grupos R de mamíferos pueden ser uno de los siguientes: (a) R=GlcNAc β (1,2); (b) R=Gal β (1,4)-GlcNAc β (1,2); (c) R=NeuAc α (2,3)-Gal β (1,4)-GlcNAc β (1,2); (d) R=NeuGc α (2,3)-Gal β (1,4)-GlcNAc β (1,2); y (e) R=Gal α (1,3)-Gal β (1,4)-GlcNAc β (1,2). Los grupos R de las plantas pueden ser uno de los siguientes: (a) R=nulo; (b) R=GlcNAc β (1,2);

(c), R= Gal β (1,3)-GlcNAc β (1,2) y (d) R=Gal β (1,4)-GlcNAc β (1,2).



Abreviaturas: Man, manosa; GlcNAc, *N*-acetilglucosamina; Xyl, xilosa; Fuc, fucosa; Gal, galactosa; NeuAc, ácido

neuramínico (ácido siálico); *, extremo reductor de cadena de azúcar que se une a asparagina.

Figura 30: muestra las especies G0, GOX y GOXF³ de glicanos N-ligados de glicoproteínas referidas en la descripción y en las reivindicaciones de la presente invención, junto con la nomenclatura alternativa usada en la presente memoria.

- 5 Figura 31: establece la secuencia parcial de ADNc (SEC ID N°: 19; secuencia codificadora establecida en la SEC ID N°: 20) correspondiente a la isoforma n°2 de la β 1,2-xilosiltransferasa (XylT) de *Lemna minor*, y la secuencia parcial de aminoácidos (SEC ID N°: 21) codificada con la misma. Los nucleótidos denotados mediante un subrayado sencillo (-) corresponden al fragmento directo de XylT dentro de la casete de expresión de ARNi diseñada para inhibir la expresión de XylT (véase la Figura 33); los nucleótidos denotados mediante un subrayado doble (=) corresponden a la secuencia espaciadora dentro de esta casete de expresión de ARNi. El fragmento inverso de XylT de la casete de expresión de ARNi es el antisentido del fragmento directo de XylT mostrado aquí.

Figura 32: establece un alineamiento de la isoforma n°1 de XylT de *Lemna minor* de SEC ID N°: 6 con la isoforma n°2 de longitud parcial de XylT de *Lemna minor* de SEC ID N°: 21.

- 15 Figura 33: establece una estrategia para diseñar un bloqueo de ARNi de gen individual de XylT de *Lemna minor* basado en la secuencia parcial de ADN correspondiente a la isoforma n°2 de XylT.

Figura 34: establece una estrategia para diseñar un bloqueo de ARNi de gen doble de FucT y XylT de *Lemna minor*, en donde la porción XylT del bloqueo de ARNi se basa en la secuencia parcial de ADN de la isoforma n°2 de XylT.

Figura 35: muestra la actividad de unión a receptor del producto mAbl derivado de CHO y derivado de SP2/0 para el Fc γ RIIIa sobre células NK humanas recién aisladas.

- 20 Figura 36: muestra la actividad de unión a receptor del producto mAbl derivado de *Lemna* natural y del producto mAbl derivado de *Lemna* transgénica correspondiente a Fc γ RIIIa sobre células NK humanas recién aisladas tomadas del Donante 1.

- 25 Figura 37: muestra la actividad de unión a receptor del producto mAbl derivado de *Lemna* natural y del producto mAbl derivado de *Lemna* transgénica correspondiente a Fc γ RIIIa sobre células NK humanas recién aisladas tomadas de los Donantes 2 y 3.

Figura 38: muestra la actividad de unión a receptor del producto mAbl derivado de Sp2/0, del producto mAbl derivado de *Lemna* natural, y del producto mAbl derivado de *Lemna* transgénica para el Fc γ RIV de ratón.

- 30 Figura 39: muestra un diagrama del vector de expresión binaria MDXA04 para el ARNi que silencia la actividad de FucT y XylT en *Lemna*. Las regiones reticuladas muestran la posición de las secuencias génicas de región variable de cadena pesada (H) y ligera (L) de anticuerpo kappa mAb1 completamente humano MDX-060 y el ARN de horquilla (ARNi) quimérico diseñado para atacar el silenciamiento de genes de *Lemna* endógenas que codifican FucT y XylT. Promotores: P1, P2 y P3; terminador: T; marcador seleccionable: SM; frontera izquierda: LB; frontera derecha, RB. El vector de expresión MDXA01 usado para expresar el mAb de MDX-060 en *Lemna* natural no contenía la región de ARN de horquilla.

- 35 Figura 40: muestra la actividad de glicosiltransferasa en las líneas de *Lemna* natural y de ARNi de MDX-060 LEX^{Opt}. Se incubaron membranas microsomaes de plantas naturales (WT) y de ARNi de MDX-060 LEX^{Opt} (se indican los números de línea) en presencia de un tampón de reacción que contenía GDP-Fuc, UDP-Xyl y GnGn-dabsil-péptido aceptor. Los picos de masas correspondientes a productos fucosilados (barras blancas) o xilosilados (barras negras) sintetizados por microsomas a partir de cada línea se midieron mediante MALDI-TOF MS de modo reflectrón positivo y se normalizaron, en porcentaje, con respecto al control WT positivo. Las membranas naturales hervidas (BWT) indican cuentas de ión de ruido de fondo.

- 40 Figura 41: muestra la SDS-PAGE de extractos de plantas y proteína A o muestras purificadas con hidroxipatita procedente de MDX-060^{Opt} en condiciones no reductoras (Figura 41A) y reductoras (Figura 41B), respectivamente. El MAb purificado procedente de una línea celular CHO (CHO MDX-060) se usó como control positivo. Se incluyeron en los geles marcadores de peso molecular Mark12. Los geles fueron teñidos con Azul Coloidal.

- 45 Figura 42: muestra los espectros obtenidos en el análisis espectrométrico de masas MALDI-TOF en modo reflectrón negativo de N-glicanos marcados con 2-AA liberados de mAbs MDX-060 expresados en CHO (CHO MDX-060), *Lemna* natural (LEX MDX-060) o *Lemna* transformada con construcción de ARNi XylT/FucT (LEX^{Opt} MDX-060). Se identificaron picos significativos a través de las masas correspondientes ([M-H]⁻). El * indica la localización de artefactos de matriz.

Figura 43: muestra los espectros obtenidos en el análisis MS NP-HPLC-QTOF de N-glicanos marcados con 2-AA liberados de mAbs MDX-060 expresados en CHO (CHO MDX-060), *Lemna* natural (LEX MDX-060) o *Lemna* transformada con la construcción de ARNi de XylT/FucT (LEX^{Opt} de MDX-060). Los N-glicanos marcados con 2-AA fueron separados mediante cromatografía de fase normal y se detectaron mediante fluorescencia. Los picos más

abundantes procedentes de cada muestra (marcados a-i) se caracterizaron mediante MS QTOF en modo negativo en línea y se muestran los correspondientes espectros de masas QTOF ($[M-2H]^{2-}$).

Figura 44: muestra la actividad in vitro de mAbs de MDX-060 medida mediante análisis de citometría de flujo de unión de mAb de CHO MDX-060, LEX o LEX^{Opt} glico-optimizados a CD30 expresado en células L540. Las células L540 se incubaron con concentraciones crecientes del anticuerpo indicado como se describe en el Ejemplo 6 incluido en la presente memoria. Se representa la intensidad de fluorescencia media geo (GMFI) frente a la concentración variable de mAb usado. ■: CHO MDX-060; ▲: LEX MDX-060; ▼: LEX^{Opt} MDX-060.

Figura 45: muestra la unión de equilibrio de mAb glico-optimizado y natural con dos alotipos de FcRγIIIa humanos diferentes (Val¹⁵⁸ o Phe¹⁵⁸). La señal de unión de FcRγIIIa se ajustó a un modelo de unión en un sitio. ■: CHO MDX-060; ▲: LEX MDX-060; ▼: LEX^{Opt} MDX-060.

Figura 46: muestra la actividad ADCC de mAb de MDX-060 derivado de CHO, LEX (glicosilación de *Lemna* natural) o LEX^{Opt} (ARNi de *Lemna* transgénica). Se incubaron células efectoras humanas procedentes de donante homocigoto FcγRIIIaPhe¹⁵⁸ y de donante heterocigoto FcγRIIIaPhe/Val¹⁵⁸ con células L540 marcadas con BATDA con una relación efector:diana de 50:1 en presencia de concentraciones crecientes de los anticuerpos indicados. Se representa el porcentaje de lisis específica para cada concentración de mAb. Se usó antígeno que no reconoce mAb 1 humano sobre células L540 como un isótopo de control en todos los experimentos. Se calcularon los valores EC₅₀ (μg/mL), las constantes de unión y el porcentaje de lisis máxima usando el software GraphPad Prism 3.0. ■: CHO MDX-060; ▲: LEX MDX-060; ▼: LEX^{Opt} MDX-060.

Figura 47: muestra el análisis de masas intactas de las composiciones de mAb de LEX MDX-060 producidas en *L. minor* natural que comprende la construcción MDXA01. Cuando las expresiones de XylT y FucT no son suprimidas en *L. minor*, la composición de mAb de LEX MDX-060 producida recombinantemente comprende al menos 7 glicoformas diferentes, siendo la glicoforma G0XF³ la especie predominante presente. Cabe destacar la ausencia de un pico que representa la glicoforma G0.

Figura 48: muestra el análisis de masas de glicano de la cadena pesada de mAb de LEX MDX-060 producida en *L. minor* natural que comprende la construcción MDXA01. Cuando las expresiones de XylT y FucT no son suprimidas en *L. minor*, la especie de *N*-glicano predominante presente es G0XF³, con picos principales adicionales que reflejan especies GOX. Cabe destacar la presencia menor de especies de glicano G0.

Figura 49: muestra el análisis de masas intactas de las composiciones de mAb de LEX^{Opt} MDX-060 producidas en *L. minor* transgénica que comprende la construcción MDXA04. Cuando las expresiones de XylT y FucT son suprimidas en *L. minor*, la composición de mAb intacta contiene solo *N*-glicanos G0. Adicionalmente, la composición es sustancialmente homogénea para la glicoforma G0 (pico 2), en donde ambos sitios de glicosilación están ocupados por especies de *N*-glicano G0, con dos picos menores que reflejan cantidades traza de glicoformas precursoras (pico 1, que muestra mAb que tiene una región Fc en la que el dominio C_{H2} de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_{H2} de otra cadena pesada está sin glicosilar; y pico 3, que muestra un mAb que tiene una región Fc en donde el sitio de glicosilación de Asn 297 de cada uno de los dominios C_{H2} tiene una especie glicano G0 unida, con una tercera especie de glicano G0 unida a un sitio adicional de glicosilación dentro de la estructura mAb).

Figura 50: muestra el análisis de masas de glicano de la cadena pesada del mAb de LEX^{Opt} MDX-060 producido en *L. minor* transgénica que comprende la construcción MDXA04. Cuando las expresiones de XylT y FucT son suprimidas en *L. minor*, la única especie de *N*-glicano fácilmente detectable unida a los sitios de glicosilación de Asn 297 de los dominios C_{H2} de las cadenas pesadas es G0.

Figuras 51A (análisis MALDI) y 51B (análisis HPLC): muestran que el perfil de glicosilación homogénea exhibido por mAbI producido en *L. minor* transgénica (línea 24) que comprende la construcción de ARNi de mAbI04 se observó consistentemente con una producción escalada. Dicho perfil de glicosilación era consistente a lo largo de un periodo de 8 meses de mantenimiento continuo de la línea transgénica a través de expansión clonal.

Figura 52: muestra que la supresión de las expresiones de FucT y XylT usando la construcción de mAbI04 ARNi químico de la Figura 12 da como resultado glicoproteínas endógenas que tienen un modelo de glicosilación homogéneo consistente con el observado para glicoproteínas recombinantes. Para esta figura, el residuo de xilosa β1,2-ligada unida a la estructura de núcleo de trimanosa se designa mediante el símbolo de estrella.

Figura 53: muestra la estructura de *N*-glicanos complejos descritos en el Ejemplo 6 presentado en esta memoria. M = manosa; Gn = *N*-acetilglucosamina; A = galactosa; X = xilosa; F = fucosa.

Descripción detallada de la invención

La presente invención se describirá ahora más detalladamente a partir de este punto en referencia a las figuras anexas, en las que se muestran algunas de las realizaciones de la invención, pero no todas. De hecho, esta invención puede realizarse en muchas formas diferentes y no debería considerarse limitada a las realizaciones establecidas en la presente memoria; más bien, estas realizaciones se proporcionan de tal modo que esta

descripción satisfará los requisitos legales aplicables. Los números similares se refieren a elementos similares.

Muchas modificaciones y otras realizaciones de la invención establecida en la presente memoria vendrán a la mente del especialista en la técnica a la que pertenece la invención, que tengan el beneficio de los conocimientos presentados en la presente descripción y en las figuras asociadas. Por tanto, debe entenderse que la invención no debe limitarse a las realizaciones específicas descritas, y que se pretende que las modificaciones y otras realizaciones estén incluidas dentro del alcance de las reivindicaciones anexas. Aunque en la presente memoria se emplean términos específicos, se usan únicamente en un sentido genérico y descriptivo y no con propósitos de limitación.

La presente invención proporciona métodos para alterar el modelo de N-glicosilación de polipéptidos homólogos y heterólogos producidos en una planta, particularmente una planta que actúa como sistema de expresión para proteínas recombinantes de interés. Los métodos comprenden el uso de construcciones de nucleótidos que comprenden una o más secuencias que son capaces de inhibir la expresión de α 1,3-fucosiltransferasa (FucT) y β 1,2-xilosiltransferasa (XylT) en una planta. También se describen composiciones para uso en la práctica de los métodos de la invención, y composiciones que comprenden glicoproteínas que tienen un perfil de glicosilación sustancialmente homólogo, que incluyen anticuerpos monoclonales optimizados con glicano con una función efectora mejorada.

Definiciones:

“Polipéptido” se refiere a cualquier proteína o péptido monomérico o multimérico.

“Polipéptido biológicamente activo” se refiere a un polipéptido que tiene la capacidad de llevar a cabo una o más funciones biológicas o un conjunto de actividades normalmente atribuidas al polipéptido en un contexto biológico. Los especialistas en la técnica apreciarán que la expresión “biológicamente activo” incluye polipéptidos en los que la actividad biológica es alterada en comparación con la proteína nativa (por ejemplo, suprimida o potenciada), siempre que la proteína tenga una actividad suficiente para ser de interés para su uso en procesos industriales o químicos o como reactivo terapéutico, vacuna o de diagnóstico. La actividad biológica se puede determinar mediante cualquier método disponible en la técnica. Por ejemplo, la actividad biológica de miembros de la familia interferón de proteínas se puede determinar mediante una serie de métodos que incluyen su interacción con anticuerpos específicos en interferón, su capacidad para aumentar la resistencia a una infección vírica, o su capacidad para modular la transcripción de dianas génicas reguladas por interferón. De un modo similar, la actividad biológica de anticuerpos monoclonales se pueden determinar mediante cualquiera de una serie de métodos que incluyen, aunque sin limitación, ensayos para medir la especificidad de unión y la función efectora, por ejemplo, usando ensayos para determinar la actividad de citotoxicidad celular dependiente de anticuerpos (ADCC) y la actividad de citotoxicidad dependiente de complemento (CDC).

Por “célula hospedante” se entiende una célula que comprende una secuencia de ácido nucleico heteróloga de la invención. Aunque las secuencias de ácido nucleico de la invención, y los fragmentos y variantes de las mismas, se pueden introducir en cualquier célula de interés, de interés particular son las células hospedante vegetales. En algunas realizaciones, las células hospedantes vegetales de una planta que actúa como hospedante para la expresión de proteínas recombinantes, por ejemplo, un sistema de expresión vegetal para la producción de proteínas de mamífero recombinante de interés tal como se indica en la presente memoria más adelante.

Con “polipéptido heterólogo de interés” se pretende indicar un polipéptido que no es expresado por células hospedantes en la naturaleza. A la inversa, un “polipéptido homólogo” pretende indicar un polipéptido que se produce de forma natural dentro de las células del hospedante. Los polipéptidos heterólogos y homólogos que sufre N-glicosilación post-traducciona se denominan en la presente memoria glicoproteínas heterólogas u homólogas. De acuerdo con los métodos de la presente invención, el modelo de N-glicosilación de ambas glicoproteínas heterólogas y homólogas es alterado dentro de células de un hospedante vegetal de tal modo que dichas glicoproteínas tienen un modelo de N-glicosilación que es más similar al observado con hospedantes mamíferos.

Los ejemplos de secuencias de nucleótidos heterólogos y de polipéptidos heterólogos de interés incluyen, aunque sin limitación, secuencias que codifican polipéptidos de mamífero, tales como la insulina, la hormona de crecimiento, α -interferón, β -interferón, β -glucocerebrosidasa, β -glucoronidasa, proteína de retinoblastoma, proteína p53, angioestatina, leptina, eritropoietina (EPO), factor estimulante de colonia de granulocito macrófago, plasminógeno, activador de plasminógeno de tejido, factores de coagulación sanguínea, por ejemplo, Factor VII, Factor VIII, Factor IX y proteína C activada, alfa 1-antitripsina, anticuerpos monoclonales (mAbs), fragmentos de Fab, anticuerpos de cadena sencilla, citoquinas, receptores, hormonas, vacunas humanas, vacunas animales, péptidos y albúmina de suero.

Para los propósitos de la presente invención, los términos “N-glicano”, “glicano N-ligado” y “glicano” se usan intercambiamente y se refieren a un oligosacárido N-ligado, por ejemplo, uno que está o estuvo unido mediante un residuo de N-acetilglucosamina (GlcNAc) ligado al nitrógeno de amida de un residuo de asparagina en una proteína. Los azúcares predominantes encontrados en las glicoproteínas son glucosa, galactosa, manosa, fucosa, N-acetilgalactosamina (GalNAc), N-acetilglucosamina (Glc-Nac), y ácido siálico (por ejemplo, ácido N-acetil-

neuramínico (NeuAc)). El procesamiento de los grupos de azúcares se produce co-traduccionalmente en el lumen del ER y continúa en el aparato de Golgi para glicoproteínas *N*-ligadas.

Con “estructura de núcleo oligomanosídico” o “estructura de núcleo de trimanosa” de un *N*-glicano complejo se pretende indicar la estructura de núcleo mostrada en la Figura 29A, en donde el núcleo comprende tres residuos de monosacárido de manosa (Man) y dos de *N*-acetilglucosamina (GlcNAc) que están unidos al residuo de asparagina de la glicoproteína. El residuo de asparagina generalmente está dentro de la secuencia de péptido conservada Asn-Xxx-Thr o Asn-Xxx-Ser, en donde Xxx es cualquier residuo excepto prolina, aspartato o glutamato. Las posteriores etapas de glicosilación producen la estructura final de *N*-glicano complejo. La estructura de núcleo de trimanosa se denota en la presente memoria como “Man₃GlcNAc₂”.

Los *N*-glicanos unidos a glicoproteínas difieren respecto al número de ramas (antenas) que comprenden azúcares periféricos (por ejemplo, GlcNAc, galactosa, fucosa y ácido siálico) que se añaden a la estructura de núcleo de trimanosa. Los *N*-glicanos se clasifican habitualmente de acuerdo a sus constituyentes ramificados (por ejemplo, complejos, de alta manosa o híbridos). Un *N*-glicano de tipo “complejo” normalmente tiene al menos un GlcNAc unido al brazo de 1,3-manosa y al menos un GlcNAc unido al brazo de 1,6-manosa del núcleo de “trimanosa”. Cuando hay un GlcNAc unido a cada brazo de manosa, la especie de *N*-glicano se denota en la presente memoria como “GlcNAc₂Man₃GlcNAc₂” o “GnGn”. Cuando solo hay un GlcNAc unido, la especie de *N*-glicano se denota en la presente memoria como “GlcAc₁Man₃GlcNAc₂”, en donde el GlcNAc está unido al brazo de 1,3-manosa (denominado “MGn” aquí) o al brazo de 1,6-manosa (denominado aquí “GnM”) (véase la Figura 30). Los *N*-glicanos complejos también pueden tener residuos de azúcar de galactosa (“Gal”) o *N*-acetilgalactosamina (“GalNAc”) que opcionalmente están modificados con ácido siálico o sus derivados (por ejemplo, “NeuAc”, donde “Neu” se refiere a ácido neuramínico y “Ac” se refiere a acetilo). Cuando hay un residuo de azúcar de galactosa unido a cada GlcNAc de cada brazo de manosa, la especie de glicano *N*-ligado se denota en la presente memoria “Gal₂GlcNAc₂Man₃GlcNAc₂”. Los *N*-glicanos complejos también pueden presentar sustituciones intracadena que comprenden “biseccionar” GlcNAc y la fucosa de núcleo (“Fuc”). Los *N*-glicanos complejos también pueden tener múltiples antenas en el “núcleo de trimanosa”, a menudo denominadas “múltiples glicanos antenarios”. Un tipo de *N*-glicano de “alta manosa” tiene cinco o más residuos de manosa. Un *N*-glicano “híbrido” tiene al menos un GlcNAc sobre el extremo del brazo de 1,3-manosa del núcleo de trimanosa y cero o más manosas en el brazo de 1,6-manosa del núcleo de trimanosa.

Las expresiones “glicano G0” y “estructura de glicano G0” y “especie de glicano G0” se usan intercambiamente y pretenden indicar el glicano *N*-ligado complejo que tiene la estructura GlcNAc₂Man₃GlcNAc₂, en donde no hay presentes ácidos siálicos (NeuAcs) terminales o residuos de galactosa (Gal) terminales. Si el glicano G0 comprende un residuo de fucosa (“Fuc”) unido a la estructura de núcleo de trimanosa, es denominado en la presente memoria “glicano G0F³” (que tiene el residuo de fucosa α1,3-ligada específico de planta) o “glicano G0F⁶” (que tiene el residuo de fucosa α1,6-ligada de mamífero). En las plantas, un glicano G0 que comprende el residuo de xilosa β1,2-ligado específico de planta unido a la estructura de núcleo de trimanosa se denomina en la presente memoria “glicano GOX”, y un glicano G0 que comprende ambos, el residuo de xilosa β1,2-ligado específico de planta y el residuo de fucosa α1,3-ligado específico de planta, unidos a la estructura de núcleo de trimanosa se denomina en la presente memoria “glicano G0XF³”.

Los términos “glicano G1” y “estructura de glicano G1” y “especies de glicano G1” se usan intercambiamente y pretenden indicar el glicano *N*-ligado complejo que tiene una estructura GlcNAc₂Man₃GlcNAc₂, en donde un residuo terminal de galactosa (Gal) está unido al brazo de 1,3-manosa o al de 1,6-manosa, y no hay ácidos siálicos terminales presentes. Los términos “glicano G2” y “estructura de glicano G2” y “especies de glicano G2” se usan intercambiamente y pretenden indicar el glicano *N*-ligado complejo que tiene la estructura GlcNAc₂Man₃GlcNAc₂, en donde hay un residuo terminal de galactosa (Gal) unido al brazo de 1,3-manosa y al brazo de 1,6-manosa, y no hay ácidos siálicos terminales presentes.

El término “glicoforma” tal como se usa en la presente memoria se refiere a una glicoproteína que contiene una estructura o estructuras de carbohidrato particular. Por tanto, por ejemplo, una “glicoforma G0” se refiere a una glicoproteína que comprende solo especies de glicano G0 unidas a sus sitios de glicosilación. Se reconoce que una glicoproteína que tiene más de un sitio de glicosilación puede tener las mismas especies de glicano unidas a cualquiera de los sitios de glicosilación, o puede tener diferentes especies de glicano unidas a diferentes sitios de glicosilación. De este modo, diferentes modelos de unión de glicano dan lugar a diferentes glicoformas de una glicoproteína.

El término “perfil de glicosilación” pretende significar la “huella dactilar” característica de la especie de *N*-glicano representativa que ha sido liberada de una composición de glicoproteína o de un producto de glicoproteína, tanto enzimática como químicamente, y a continuación ha sido analizada para determinar su estructura de carbohidrato, por ejemplo, usando LC-HPLC o MS MALDI-TOF, y otros similares. Véase, por ejemplo, la revisión publicada en *Current Analytical Chemistry*, Vol. 1, N° 1 (2005), pág. 28-57.

Las expresiones “sustancialmente homogénea”, “sustancialmente uniforme” y “homogeneidad sustancial” en el contexto de un perfil de glicosilación para una composición de glicoproteína o un producto de glicoproteína se usan intercambiamente y pretenden indicar un perfil de glicosilación en el que al menos un 90%, un 91%, un 92%, un

93%, un 94%, un 95%, un 96%, un 97%, un 98% o al menos un 99% de las especies totales de *N*-glicano del perfil están representadas por una especie de *N*-glicano deseada, con una cantidad traza de especies de *N*-glicano precursoras que aparecen en el perfil. Por "cantidad traza" se entiende que cualquier especie dada de *N*-glicano precursora que esté presente en el perfil de glicosilación está presente en menos de un 5%, preferiblemente en menos de un 4%, menos de un 3%, menos de un 2%, menos de un 1% e incluso menos de un 0,5% o incluso menos de un 0,1% de la cantidad total de especies de *N*-glicano que aparecen en el perfil. Con especies de *N*-glicano "precursoras" se pretende indicar una especie de *N*-glicano que está procesada de manera incompleta. Los ejemplos de especies de *N*-glicano precursoras presentes en cantidades traza en las composiciones de glicoproteína o en los productos de glicoproteína de la invención, y que por tanto aparecen en sus perfiles de glicosilación, son las especies de *N*-glicano precursoras descritas anteriormente, $\text{Man}_3\text{GlcNAc}_2$, MGn ($\text{GlcNAc1Man}_3\text{GlcNAc}_2$, en donde GlcNAc1 está unida al brazo de 1,3-manosa), y GnM ($\text{GlcNAc1Man}_3\text{GlcNAc}_2$, en donde GlcNAc1 está unida al brazo de 1,6 manosa).

En la presente memoria, la especie de *N*-glicano deseada dentro de un producto o composición de glicoproteínas es G0, de tal modo que un perfil de glicosilación sustancialmente homogéneo para dicho producto o composición es uno en el que al menos un 90%, un 91%, un 92%, un 93%, un 94%, un 95%, un 96%, un 97%, un 98% o al menos un 99% de las especies totales de *N*-glicano que aparecen en el perfil de glicosilación correspondiente al producto o composición están representadas por especies de glicano G0, con una cantidad traza de especies de *N*-glicano precursoras que aparezcan en el perfil de glicosilación. Para dicha composición, una especie de *N*-glicano precursora representativa que aparece en el perfil de glicosilación serían las especies de *N*-glicano precursoras descritas anteriormente, $\text{Man}_3\text{GlcNAc}_2$, MGn ($\text{GlcNAc1Man}_3\text{GlcNAc}_2$, en donde GlcNAc1 está unida al brazo de 1,3-manosa), y GnM ($\text{GlcNAc1Man}_3\text{GlcNAc}_2$, en donde GlcNAc1 está unida al brazo de 1,6 manosa).

Las expresiones "sustancialmente homogéneo", "sustancialmente uniforme" y "homogeneidad sustancial" en el contexto de una composición de glicoproteína o de un producto de glicoproteína se usan intercambiamente y pretenden indicar un producto de glicoproteína o una composición de glicoproteína en la que al menos un 90%, un 91%, un 92%, un 93%, un 94%, un 95%, un 96%, un 97%, un 98% o al menos un 99% de la glicoproteína presente en el producto o composición está representado por una glicofoma deseada, con una cantidad traza de precursor o de glicofomas no deseadas presentes en la composición. Por "cantidad traza" se pretende indicar que cualquier precursor o glicofoma no deseada dados que esté presente en el producto o composición de glicoproteína lo está en menos de un 5%, preferiblemente en menos de un 4%, menos de un 3%, menos de un 2%, menos de un 1% e incluso menos de un 0,5% o incluso menos de un 0,1% de la glicoproteína total. Con glicofoma "precursora" se pretende indicar una glicofoma en la que al menos un sitio de glicosilación está sin glicosilar, o está ocupado por una especie de *N*-glicano que representa un precursor de la especie de *N*-glicano deseada, o una glicofoma en donde hay presentes uno o más sitios de glicosilación adicionales, respecto a la glicofoma deseada, y está ocupado (es decir, tiene unido) por la especie de *N*-glicano deseada o por una especie de *N*-glicano no deseada.

Por tanto, una composición o producto de glicoproteína sustancialmente homogéneo que comprende la glicofoma G0 es una composición o producto en la que al menos un 90%, un 91%, un 92%, un 93%, un 94%, un 95%, un 96%, un 97%, un 98% o al menos un 99% de la glicoproteína presente en el producto o composición está representado por la glicofoma G0, en donde todos los sitios de glicosilación anticipados está ocupados por especies de glicano G0, estando presentes una cantidad traza de glicofomas precursoras o no deseadas en la composición. En dicha composición, una glicofoma precursora representativa sería una en la que los sitios de glicosilación no están ocupados, y un ejemplo de glicofoma no deseada sería una glicofoma que tiene una mezcla de glicano G0 y especies de glicano G0X ó G0XF³ unidas a sus sitios de glicosilación.

El término "anticuerpo" se usa en el sentido más amplio y cubre completamente anticuerpos ensamblados, fragmentos de anticuerpos que se pueden unir a antígeno (por ejemplo, Fab', F'(ab)₂, Fv, anticuerpos de cadena sencilla, diacuerpos), y péptidos recombinantes que comprenden los anteriores. Los anticuerpos representan una de muchas glicoproteínas contempladas por los métodos y composiciones de la presente invención. También se contemplan derivados de anticuerpos en la presente invención. Los derivados incluyen proteínas de fusión que comprenden una inmunoglobulina o porción de la misma, tal como una región Fc que tiene un dominio C_H2.

El término "anticuerpo monoclonal" tal como se usa en la presente memoria se refiere a un anticuerpo obtenido a partir de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que comprenden la población son idénticos excepto por las posibles mutaciones naturales que puedan darse en cantidades menores.

Los "anticuerpos nativos" y las "inmunoglobulinas nativas" habitualmente son glicoproteínas heterotetraméricas de aproximadamente 150.000 daltons, compuestas por dos cadenas ligeras (L) idénticas y dos cadenas pesadas (H) idénticas. Cada cadena ligera está unida a una cadena pesada mediante un enlace covalente de disulfuro, mientras que el número de enlaces de disulfuro varía entre las cadenas pesadas de los diferentes isotipos de inmunoglobulina. Cada cadena pesada y ligera también tiene enlaces de disulfuro intracadena espaciados regularmente. Cada cadena pesada tiene un extremo y un dominio variable (V_H) seguido de un número de dominios constantes. Cada cadena ligera tiene un dominio variable en un extremo (V_L) y un dominio constante en su otro extremo; el dominio constante de la cadena ligera se alinea con el primer dominio constante de la cadena pesada, y el dominio variable de la cadena ligera se alinea con el dominio variable de la cadena pesada. Se cree que residuos

de aminoácido particulares forman una interfase entre los dominios variables de cadena ligera y cadena pesada.

El término "variable" se refiere al hecho de que determinadas porciones de los dominios variables difieren extensivamente en secuencia entre los anticuerpos y se usan en la unión y especificidad de cada anticuerpo concreto para su antígeno particular. Sin embargo, la variabilidad no se distribuye uniformemente a lo largo de los dominios variables de los anticuerpos. Está concentrada en tres segmentos denominados regiones de determinación de la complementariedad (CDRs) o regiones hipervariables, tanto en los dominios variables de cadena ligera como en los de cadena pesada. Las porciones más altamente conservadas de dominios variables se denominan regiones estructurales (FR). Los dominios variables de cadenas nativas pesadas y ligeras comprenden cuatro regiones FR, adoptando mayormente una configuración de lámina β , conectada por tres CDRs, que forma lazos que conectan, y en algunos casos que forman parte de la estructura de lámina β . Las CDRs de cada cadena se mantienen unidos en estrecha proximidad mediante regiones FR y, con las CDRs de la otra cadena, contribuyen a la formación del sitio de unión a antígeno de los anticuerpos (véase Kabat et al. (1991) NIH Publ. N° 91-3242, Vol. I, páginas 647-669).

Los dominios constantes no están implicados directamente en la unión de un anticuerpo con un antígeno, pero exhiben varias funciones efectoras, tales como la unión de receptor Fc (FcR), la participación del anticuerpo en la citotoxicidad celular dependiente de anticuerpo (ADCC), la opsonización, la iniciación de citotoxicidad dependiente de complemento (actividad CDC), y la desgranulación de mastocitos.

"Fragmentos de anticuerpo" comprende una porción de un anticuerpo intacto, preferiblemente la región de unión a antígeno o la región variable del anticuerpo intacto. Los ejemplos de fragmentos de anticuerpo incluyen Fab, Fab', F(ab')₂ y fragmentos Fv; diacuerpos; anticuerpos lineales (Zapata et al. (1995) Protein Eng. 8(10): 1057-1062); moléculas de anticuerpo de cadena sencilla; y anticuerpos multiespecíficos formados a partir de fragmentos de anticuerpo. La digestión con papaína de anticuerpos produce dos fragmentos de unión a antígeno idénticos, denominados fragmentos "Fab", cada uno con un sitio de unión a antígeno individual, y un fragmento "Fc" residual, cuyo nombre refleja su capacidad para cristalizar fácilmente. El tratamiento con pepsina produce un fragmento F(ab')₂ que tiene dos sitios de combinación de antígeno y todavía es capaz de reticular antígeno.

"Fv" es el fragmento de anticuerpo mínimo que contiene un sitio de reconocimiento y unión a antígeno completo. En una especie Fv de dos cadenas, dicha región consta de un dímero de un dominio variable de cadena pesada y uno de cadena ligera en estrecha asociación, no covalente. En una especie Fv de cadena sencilla, un dominio variable de cadena pesada y uno de cadena ligera pueden ligarse covalentemente mediante un ligando de péptido flexible de tal modo que las cadenas ligera y pesada se pueden asociar en una estructura "dimérica" análoga a la presente en una especie Fv de dos cadenas. Es en esta configuración en la que las tres CDRs de cada dominio variable interaccionan para definir un sitio de unión de antígeno sobre la superficie del dímero V_H-V_L. Colectivamente, las seis CDRs confieren especificidad de unión a antígeno al anticuerpo. Sin embargo, incluso un único dominio variable (o la mitad de un Fv que comprende solo tres CDRs específicas de un antígeno) tiene la capacidad de reconocer y unirse a antígeno, aunque con una menor afinidad que el sitio de unión completo.

El fragmento Fab también contiene el dominio constante de la cadena ligera y el primer dominio constante (C_{H1}) de la cadena pesada. Los fragmentos Fab difieren de los fragmentos Fab' en la adición de nuevos residuos en los extremos carboxi del dominio de C_{H1} de cadena pesada que incluye una o más cisteínas de la región bisagra del anticuerpo. Fab'-SH es la designación usada en la presente memoria para el Fab' en el que el(los) residuo(s) de cisteína de los dominios constantes portan un grupo tiol libre. Los fragmentos de anticuerpo F(ab')₂ se produjeron originalmente como pares de fragmentos Fab' que tienen cisteínas de bisagra entre ellos. También se conocen otros acoplamiento químicos de fragmentos de anticuerpo.

Las "cadenas ligeras" de anticuerpos (inmunoglobulinas) de cualquier especie vertebrada se pueden asignar a uno de dos tipos claramente distintos, denominados kappa (κ) y lambda (λ), en base a las secuencias de aminoácido de sus dominios constantes.

Dependiendo de la secuencia de aminoácidos del dominio constante de sus cadenas pesadas, las inmunoglobulinas se pueden asignar a clases diferentes. Existen cinco clases principales de inmunoglobulinas humanas: IgA, IgD, IgE, IgG e IgM, y varias de éstas pueden dividirse a su vez en subclases (isotipos), por ejemplo, IgG1, IgG2, IgG3, IgG4, IgA e IgA2. Los dominios constantes de cadena pesada que corresponden a las diferentes clases de inmunoglobulinas se denominan alfa, delta, épsilon, gamma y mu, respectivamente. Las estructuras de subunidades y las configuraciones tridimensionales de las diferentes clases de inmunoglobulinas son bien conocidas. Los diferentes isotipos tienen diferentes funciones efectoras. Por ejemplo, los isotipos de IgG 1 e IgG3 humanas median en la actividad de citotoxicidad mediada por célula dependiente de anticuerpo (ADCC).

Las inmunoglobulinas tienen una glicosilación N-ligada conservada de la región Fc de cada una de las dos cadenas pesadas. Por tanto, por ejemplo, las inmunoglobulinas del tipo IgG tienen dominios C_{H2} glicosilados que portan oligosacáridos N-ligados en la asparagina 297 (Asn-297). Existen diferentes glicoformas de inmunoglobulinas dependiendo de las especies de N-glicano particulares unidas a cada uno de estos dos sitios de glicosilación, y dependiendo del grado al cual ambos sitios son glicosilados dentro de una composición de inmunoglobulina.

"Secuencia de nucleótidos de interés", tal como se usa en la presente memoria en referencia a la expresión de

polipéptidos heterólogos se refiere a cualquier secuencia de polinucleótidos que codifica un polipéptido heterólogo destinado a la expresión en un hospedante, particularmente un hospedante vegetal, por ejemplo, en una planta superior, que incluye los miembros de las *dicotyledonaceae* y *monocotyledonaceae*. Por ejemplo, las secuencias de polinucleótidos que codifican polipéptidos terapéuticos (por ejemplo, para usos veterinarios o médicos) o inmunogénicos (por ejemplo, para vacunación) se pueden expresar usando hospedantes vegetales transformados, por ejemplo, lenteja de agua, de acuerdo con la presente invención.

El uso del término “polinucleótido” no pretende limitar la presente invención a polinucleótidos que comprenden ADN. Los especialistas en la técnica reconocerán que los polinucleótidos pueden comprender ribonucleótidos y combinaciones de ribonucleótidos y desoxirribonucleótidos. Dichos desoxirribonucleótidos y ribonucleótidos incluyen tanto moléculas naturales como análogos sintéticos. Los polinucleótidos de la invención también abarcan todas las formas de secuencias que incluyen, aunque sin limitación, formas de cadena sencilla, formas de cadena doble, horquillas, estructuras de tallo-y-lazo, y otras similares.

Los términos “inhibir”, “inhibición” e “que inhibe”, tal como se usan en la presente memoria, se refieren a cualquier disminución de la expresión o de la función de un producto génico diana, que incluye una reducción relativa de la expresión o de la función hasta la completa eliminación de la expresión o función del producto génico diana. El término “expresión” tal como se usa en la presente memoria en el contexto de un producto génico se refiere a la biosíntesis de dicho producto génico, que incluye la transcripción y/o la traducción y/o el ensamblaje del producto génico. La inhibición de la expresión o de la función de un producto génico diana (es decir, un producto génico de interés) puede ser en el contexto de una comparación entre dos plantas cualesquiera, por ejemplo, la expresión o la función de un producto génico diana de una planta alterada genéticamente frente a la expresión o función de dicho producto génico diana en la planta natural correspondiente. Alternativamente, la inhibición de la expresión o de la función del producto génico diana puede ser en el contexto de una comparación entre células vegetales, orgánulos, órganos, tejidos o partes de plantas dentro de la misma planta o entre diferentes plantas, e incluye comparaciones entre etapas de desarrollo o temporales dentro de la misma planta o entre diferentes plantas. Se puede usar cualquier método o composición que regula a la baja la expresión de un producto génico diana, tanto a nivel de transcripción como de traducción, o que regula a la baja la actividad funcional del producto génico diana, para lograr la inhibición de la expresión o de la función del producto génico diana.

El término “secuencia inhibidora” abarca cualquier secuencia de polinucleótido o de polipéptido que sea capaz de inhibir la expresión de un producto génico diana, por ejemplo, a nivel de transcripción o traducción, o que sea capaz de inhibir la función de un producto génico diana. Los ejemplos de secuencias inhibidoras incluyen, aunque sin limitación, secuencias de polinucleótidos o de polipéptidos de longitud completa, secuencias de polinucleótidos o de polipéptidos truncadas, fragmentos de secuencias de polinucleótidos o de polipéptidos, variantes de secuencias de polinucleótidos o de polipéptidos, secuencias de nucleótidos con orientación sentido, secuencias de nucleótidos con orientación antisentido, el complemento de una secuencia de nucleótidos con orientación sentido o antisentido, regiones invertidas de secuencias de nucleótidos, horquillas de secuencias de nucleótidos, secuencias de nucleótidos de cadena doble, secuencias de nucleótidos de cadena sencilla, combinaciones de las mismas, y otras similares. La expresión “secuencia de polinucleótidos” incluye secuencias de ARN, ADN, ácidos nucleicos modificados químicamente, análogos de ácido nucleico, combinaciones de las mismas, y otras similares.

Se reconoce que los polinucleótidos inhibidores incluyen secuencias de nucleótidos que inhiben la expresión de un producto génico diana directamente (es decir, no requieren transcripción) o indirectamente (es decir, requieren transcripción o transcripción y traducción). Por ejemplo, un polinucleótido inhibidor puede comprender una secuencia de nucleótidos que se sintetiza químicamente o ARN interferente pequeño (ARNsi) producido *in vitro* o micro ARN (miARN) que, cuando se introduce en una célula, tejido u órgano vegetal, silenciaría directamente, aunque de manera temporal, la expresión del producto génico de interés. Alternativamente, un polinucleótido inhibidor puede comprender una secuencia de nucleótidos que codifica una molécula de nucleótido inhibidora que se diseña para silenciar la expresión del producto génico de interés, tal como ARN de orientación sentido, ARN antisentido, ARN de cadena doble (ARNds), ARN de horquilla (ARNhp), ARNhp que contiene intrón, ARN catalítico, miARN, y otros similares. En otras realizaciones adicionales, el polinucleótido inhibidor puede comprender una secuencia de nucleótidos que codifica un ARNm, cuya traducción da lugar a un polipéptido que inhibe la expresión o la función del producto génico diana de interés. De este modo, cuando el polinucleótido inhibidor comprende una secuencia de nucleótidos que codifica una molécula de nucleótido inhibidora o un ARNm para un polipéptido, la secuencia codificadora está ligada operativamente a un promotor que conduce la expresión en un célula vegetal de tal modo que la molécula de nucleótido inhibidora codificada o el ARNm se pueden expresar.

Las secuencias inhibidoras se designan en la presente memoria con el nombre del producto génico diana. Por tanto, por ejemplo, una “secuencia inhibidora de α 1,3-fucosiltransferasa (FucT)” se referiría a una secuencia inhibidora que es capaz de inhibir la expresión de una FucT, por ejemplo, a nivel de transcripción y/o traducción, o que es capaz de inhibir la función de una FucT. De forma similar, una “secuencia inhibidora de β 1,2-xilosiltransferasa (XylT)” se referiría a una secuencia inhibidora que es capaz de inhibir la expresión de una XylT, a nivel de transcripción y/o de traducción, o que es capaz de inhibir la función de una XylT. Cuando se usa la expresión “capaz de inhibir” en el contexto de una secuencia inhibidora de polinucleótido, se pretende indicar que la propia secuencia inhibidora ejerce el efecto inhibidor; o, cuando la secuencia inhibidora codifica una molécula de nucleótido inhibidora (por ejemplo, polinucleótidos de ARN de horquilla, miARN o ARN de cadena doble), o codifica un polipéptido inhibidor (es decir, un

polipéptido que inhibe la expresión o la función del producto génico diana), después de su transcripción (por ejemplo, en el caso de una secuencia inhibidora que codifica un polinucleótido de ARN de horquilla, miARN o ARN de cadena doble) o de su transcripción y traducción (en el caso de una secuencia inhibidora que codifica un polipéptido inhibidor), el producto transcrito o traducido, respectivamente, ejerce el efecto inhibidor sobre el producto génico diana (es decir, inhibe la expresión o la función del producto génico diana).

El término “introducir” en el contexto de un polinucleótido, por ejemplo, una construcción de nucleótidos de interés, pretende significar presentar a la planta el polinucleótido de un modo tal que el polinucleótido gane acceso al interior de una célula de la planta. Cuando se va a introducir más de un polinucleótido, dichos polinucleótidos se pueden ensamblar como parte de una única construcción de nucleótido, o como construcciones de nucleótido separadas, y puede localizarse sobre el mismo o diferentes vectores de transformación. Por consiguiente, dichos polinucleótidos se pueden introducir en la célula hospedante de interés en un único evento de transformación, en eventos de transformación separados o, por ejemplo, en plantas, como parte de un protocolo de cultivo. Los métodos de la invención no dependen de un método particular para introducir uno o más polinucleótidos en una planta, solo que el(los) polinucleótido(s) gana(n) acceso al interior de al menos una célula de la planta. Los métodos para introducir polinucleótidos en plantas son conocidos en la técnica, e incluyen, aunque sin limitación, métodos de transformación temporal, métodos de transformación estable y métodos mediados por virus.

“Transformación temporal” en el contexto de un polinucleótido pretende significar que un polinucleótido se introduce en la planta y no se integra en el genoma de la planta.

Con “introducir establemente” o “introducido establemente” en el contexto de un polinucleótido introducido en una planta pretende que el polinucleótido introducido se incorpora establemente al genoma de la planta, y por tanto la planta se transforma de manera estable con el polinucleótido.

“Transformación estable” o “transformado establemente” pretende significar que un polinucleótido, por ejemplo, una construcción de nucleótido descrita en la presente memoria, introducida en una planta se integra en el genoma de la planta y es capaz de ser heredada por la progenie de la misma, más particularmente, por la progenie de múltiples generaciones sucesivas. En algunas realizaciones, las generaciones sucesivas incluyen la progenie producida a través vía reproducción sexual. Un hospedante vegetal superior que es “transformado establemente” con al menos una construcción de nucleótido que es capaz de inhibir la expresión de una FucT y/o una XylT tal como se describe en la presente memoria se refiere a un hospedante vegetal superior que tiene la(s) construcción(es) de nucleótido integrada(s) en su genoma, y es capaz de producir progenie, bien por reproducción asexual o bien por reproducción sexual, que también comprende la(s) construcción(es) de nucleótido inhibidora(s) integrada(s) establemente en su genoma, y por tanto la progenie también presentará el fenotipo deseado de tener un modelo de *N*-glicosilación alterado que se caracteriza por una reducción en la unión de residuos de α 1,3-fucosa y/o β 1,2-xilosa a los *N*-glicanos de glicoproteínas homólogos y heterólogos producidas.

Tal como se usa en la presente memoria, el término “planta” incluye la referencia a plantas enteras, órganos de plantas (por ejemplo, hojas, tallos, raíces, etc.), semillas, células vegetales y su progenie. Las partes de plantas transgénicas deben entenderse dentro del alcance la invención que comprenden, por ejemplo, células vegetales, protoplastos, tejidos, callos, embriones así como flores, óvulos, tallos, frutos, hojas, raíces, puntas de raíces, y otros similares que se originan en plantas transgénicas o en su progenie transformada previamente con una molécula de ADN de la invención, y por lo tanto que consiste al menos en parte en sus células transgénicas. Tal como se usa en la presente memoria, el término “célula vegetal” incluye, sin limitación, células de semillas, embriones, regiones meristemáticas, tejido calloso, hojas, raíces, brotes, gametofitos, esporofitos, polen y microesporas.

La clase de plantas que puede usarse en los métodos de la invención generalmente es tan amplia como las clases de plantas superiores susceptibles de someterse a técnicas de transformación, incluyendo tanto plantas monocotiledóneas (monocot) como dicotiledóneas (dicot). Los ejemplos de dicots incluyen, aunque sin limitación, legumbres que incluyen soja y alfalfa, tabaco, patatas, tomates y otros similares. Los ejemplos de monocots incluyen, aunque sin limitación, maíz, arroz, avena, cebada, trigo, miembros de la familia de las lentejas de agua, hierbas, y otros similares. “Plantas de orden inferior” se refiere a plantas que no florecen que incluyen helechos, colas de caballo, musgos, hierbas para infusión liverwort, hornwort, algas, por ejemplo, algas rojas, marrones y verdes, gametofitos, esporofitos de pteridofites, y otros similares. En algunas realizaciones, la planta de interés es un miembro de la familia de plantas de lentejas de agua.

El término “lentejas de agua” se refiere a los miembros de la familia *Lemnaceae*. Esta familia está dividida actualmente en cinco géneros y 38 especies de lentejas de agua como se indica a continuación: género *Lemna* (*L. aequinoctialis*, *L. disperma*, *L. ecuadoriensis*, *L. gibba*, *L. japonica*, *L. minor*, *L. miniscula*, *L. obscura*, *L. perpusilla*, *L. tenera*, *L. trisulca*, *L. turionifera*, *L. valdiviana*); género *Spirodela* (*S. intermedia*, *S. polyrrhiza*, *S. punctata*); género *Wolffia* (*Wa. angusta*, *Wa. arrhiza*, *Wa. australina*, *Wa. borealis*, *Wa. brasiliensis*, *Wa. columbiana*, *Wa. elongata*, *Wa. globosa*, *Wa. microscopica*, *Wa. neglecta*); género *Wolffiella* (*Wl. caudata*, *Wl. denticulata*, *Wl. gladiata*, *Wl. hyalina*, *Wl. lingulata*, *Wl. repunda*, *Wl. rotunda* y *Wl. neotropica*) y género *Landoltia* (*L. punctata*). Cualesquier otros géneros o especies de *Lemnaceae*, si existen, también son aspectos de la presente invención. La especie *Lemna* se puede clasificar usando el esquema taxonómico descrito por Landolt (1986) “Biosystematic Investigation on the

Family of Duckweeds: The family of Lemnaceae – A Monograph Study” (Geobotanischen Institut ETH, Stiftung Rubel, Zürich).

El término “nódulo de lenteja de agua” tal como se usa en la presente memoria se refiere a tejido de lenteja de agua que comprende células de lenteja de agua en donde al menos aproximadamente 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% ó 100% de las células son células diferenciadas. Una “célula diferenciada”, tal como se usa en la presente memoria, es una célula con al menos una característica fenotípica (por ejemplo, una morfología celular distintiva o la expresión de un ácido nucleico o proteína marcadora) que la distingue de células no diferenciadas o de células que se encuentran en otros tipos de tejidos. Las células diferenciadas del cultivo de nódulo de lenteja de agua que han comenzado a organizarse en frond primordium distribuido por todo el tejido. La superficie del tejido del cultivo de nódulo tiene células epidermales conectadas a cada vía plasmadesmata. Los miembros de la familia de las lentejas de agua se reproducen mediante propagación clonal, y por tanto son representativas de plantas que se propagan clonalmente.

“Codones preferidos de lentejas de agua” tal como se usa en la presente memoria se refiere a codones que tienen una frecuencia de uso de codón en lentejas de agua superior al 17%.

“Codones preferidos de *Lemna*” tal como se usa en la presente memoria se refiere a codones que tienen una frecuencia de uso de codón en el género *Lemna* superior al 17%.

“Codones preferidos de *Lemna gibba*” tal como se usan en la presente memoria se refiere a codones que tienen una frecuencia de uso de codón en *Lemna gibba* superior al 17%, en donde la frecuencia de uso de codón de *Lemna gibba* se obtuvo a partir de la Base de Datos de Uso de Codón (GenBank Release 113,0; en [http://www.kazusa.or.jp/codon/egibin/showcodon.cgi?species=Lemna+gibba+\[gbpln\]](http://www.kazusa.or.jp/codon/egibin/showcodon.cgi?species=Lemna+gibba+[gbpln])).

“Codón de inicio de traducción” se refiere al codón que inicia la traducción del ARNm transcrito desde la secuencia de nucleótidos de interés.

“Secuencia de nucleótido de contexto de iniciación de la traducción” tal como se usa en la presente memoria se refiere a la identidad de los tres nucleótidos directamente a 5' del codón de iniciación de la traducción.

“Secreción”, tal como se usa en la presente memoria, se refiere a la traslocalización de un polipéptido a través de la membrana plasmática y la pared celular de una célula de planta hospedante.

“Ligado operativamente” tal como se usa en la presente memoria en referencia a secuencias de nucleótido se refiere a múltiples secuencias de nucleótido que se colocan en una relación funcional una con otra. Generalmente, las secuencias de ADN ligadas operativamente están contiguas y, cuando es necesario para unir dos regiones codificadoras de proteína, en marco de lectura.

Polinucleótidos y Polipéptidos Aislados:

La presente invención describe polinucleótidos y polipéptidos aislados que están implicados en una modificación adicional de glicanos *N*-ligados de planta (también denominados “*N*-glicanos”), particularmente una α 1,3-fucosiltransferasa (FucT) y una β 1,2-xilosiltransferasa (XylT) identificadas en *Lemna minor*, un miembro de la familia de las lentejas de agua, y las variantes y fragmentos de dichos polinucleótidos y polipéptidos. La inhibición de la expresión de una de estas proteínas, o de ambas, o de las variantes biológicamente activas de las mismas, en una planta que expresa dichas proteínas da lugar beneficiosamente a un modelo de *N*-glicosilación que presenta una reducción en la unión de residuos de α 1,3-fucosa y β 1,2-xilosa a *N*-glicanos de glicoproteínas. En algunas realizaciones de la invención, los métodos descritos en la presente memoria proporcionan una inhibición completa de la expresión de FucT y XylT, dando lugar a un modelo de *N*-glicosilación de glicoproteínas producidas en una planta en el que los glicanos *N*-ligados están desprovistos de residuos de α 1,3-fucosa y β 1,2-xilosa.

La secuencia de ADNc de longitud completa, que incluye 5'- y 3'-UTR, correspondiente a la alfa 1,3 fucosiltransferasa (FucT) de *L. minor* se establece en la Figura 1; véase también la SEC ID N°: 1 (marco de lectura abierto establecido en la SEC ID N°: 2). La secuencia de aminoácidos predicha codificada por ella se establece en la SEC ID N°: 3. Se han identificado al menos dos isoformas del gen de FucT de *L. minor*; la homología entre las isoformas es de aproximadamente un 90%. La proteína codificada comparte alguna similitud con otras FucTs de otras plantas superiores. Véase la Figura 2. Por ejemplo, la secuencia de FucT de *L. minor* comparte aproximadamente un 50,1% de identidad de secuencia con la FucT de *Arabidopsis thaliana* mostrada en la Figura 2.

La secuencia de ADNc de longitud completa, que incluye 5'- y 3'-UTR, correspondiente a la β 1,2-xilosiltransferasa (XylT) de *L. minor* (isoforma n°1) se establece en la Figura 3; véase también la SEC ID N°: 4 (ORF establecido en la SEC ID N°: 5). La secuencia de aminoácidos predicha codificada por ella se establece en la SEC ID N°: 6. Se han identificado al menos dos isoformas del gen de XylT de *L. minor*; la homología entre las isoformas es de aproximadamente un 90%. La proteína codificada comparte alguna similitud con otras XylTs de otras plantas superiores. Véase la Figura 4. Por ejemplo, la XylT de *L. minor* comparte aproximadamente un 56,4% de identidad de secuencia con la XylT de *Arabidopsis thaliana* mostrada en la Figura 4. Una secuencia de ADNc de longitud parcial, que incluye 3'-UTR, correspondiente a la β 1,2-xilosiltransferasa (XylT) de *L. minor* (isoforma n°2) se

establece en la Figura 31; véase también la SEC ID N°: 19 (ORF establecido en la SEC ID N°: 20). La secuencia de aminoácidos predicha codificada por ella se establece en la SEC ID N°: 21. La isoforma n°2 de XylT de longitud parcial comparte una elevada identidad de secuencia con la correspondiente región de la isoforma n°1 de XylT de longitud completa, como puede observarse en el alineamiento mostrado en la Figura 32.

La invención describe composiciones de polinucleótidos o proteínas aislados o sustancialmente purificados. Un polinucleótido o proteína "aislado" o "purificado", o una porción biológicamente activa del mismo, está sustancial o esencialmente libre de los componentes que acompañan o interaccionan normalmente con el polinucleótido o proteína según se encuentra en entorno natural. Por tanto, un polinucleótido o proteína aislado o purificado está sustancialmente libre de otros materiales celulares, o de medio de cultivo cuando se produce mediante técnicas recombinantes, o está sustancialmente libre de precursores químicos u otros agentes químicos cuando se sintetiza químicamente. De forma óptima, un polinucleótido "aislado" está libre de las secuencias (de manera óptima de secuencias codificadoras de proteína) que de forma natural flanquean al polinucleótido (es decir, secuencias localizadas en los extremos 5' y 3' del polinucleótido) en el ADN genómico del organismo del cual deriva el polinucleótido. Por ejemplo, en varias realizaciones, el polinucleótido aislado puede contener menos de aproximadamente 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ó 0,1 kb de secuencia de nucleótidos que flanquean de manera natural al polinucleótido en el ADN genómico de la célula de la cual deriva el polinucleótido. Una proteína que está sustancialmente libre de material celular incluye preparaciones de proteínas que tienen menos de aproximadamente 30%, 20%, 10%, 5% ó 1% (en peso seco) de proteína contaminante. Cuando la proteína o la porción biológicamente activa de la misma se produce recombinantemente, el medio de cultivo óptimamente representa menos de aproximadamente el 30%, 20%, 10%, 5% ó 1% (en peso seco) de los precursores químicos o de productos químicos que no son de interés proteínico.

La secuencia codificadora correspondiente al gen *FucT* de *L. minor* se establece como los nucleótidos (nt) 243-1715 de la SEC ID N°: 1 y como la SEC ID N°: 2, y la secuencia de aminoácidos correspondiente al polipéptido de *FucT* codificado se establece en la SEC ID N°: 3. La secuencia codificadora correspondiente al gen de la isoforma n°1 de *XylT* de *L. minor* se establece como los nucleótidos 63-1592 de la SEC ID N°: 4 y como la SEC ID N°: 5, y la secuencia de aminoácidos correspondiente al polipéptido de *XylT* codificado se establece en la SEC ID N°: 6. La secuencia codificadora correspondiente al gen de la isoforma n°2 de *XylT* de *L. minor* de longitud parcial se establece como los nucleótidos 1-1276 de la SEC ID N°: 19 y como la SEC ID N°: 20, y la secuencia de aminoácidos correspondiente al polipéptido de *XylT* de longitud parcial se establece como la SEC ID N°: 21.

En particular, la presente invención describe polinucleótidos aislados que comprenden secuencias de nucleótido que codifican las secuencias de aminoácido mostradas en las SEC ID N°: 3, 6 y 21. Además se describen polipéptidos que tienen una secuencia de aminoácidos codificada por un polinucleótido descrito en la presente memoria, por ejemplo aquellos establecidos en las SEC ID N°: 1, 2, 4, 5, 19 y 20, y fragmentos y variantes de los mismos. También se describen las moléculas de ácido nucleico que comprenden los complementos de dichas secuencias de nucleótidos. Se reconoce que la secuencia codificadora correspondiente al gen de *FucT* y/o de *XylT* se puede expresar en una planta para sobreexpresión de la *FucT* y/o la *XylT* codificada. Sin embargo, para los propósitos de suprimir o inhibir la expresión de estas proteínas, se usarán las respectivas secuencias de nucleótidos de las SEC ID N°: 1, 2, 4, 5, 19 y 20 para diseñar construcciones para la supresión de la expresión de la respectiva proteína *FucT* y/o *XylT*. Por tanto, los polinucleótidos, en el contexto de suprimir la proteína *FucT*, se refieren a las secuencias y a los polinucleótidos que codifican *FucT* que cuando se expresan suprimen o inhiben la expresión del gen de *FucT*, por ejemplo, a través de la supresión directa o indirecta según se indica más adelante en la presente memoria. De forma similar, los polinucleótidos, en el contexto de suprimir o inhibir la proteína *XylT*, se refieren a las secuencias y polinucleótidos que codifican *XylT* que cuando se expresan suprimen o inhiben la expresión del gen de *XylT*, por ejemplo, a través de la supresión directa o indirecta según se indica más adelante en la presente memoria.

Los fragmentos y variantes de los polinucleótidos descritos y las proteínas codificadas con ellos también se describen en la presente memoria. Por "fragmento" se entiende una porción del polinucleótido de *FucT* y/o *XylT* o una porción de la secuencia de aminoácidos de *FucT* y/o *XylT* codificada con él. Los fragmentos de un polinucleótido pueden codificar fragmentos de proteína que retienen la actividad biológica de la proteína nativa y que por tanto tienen actividad de *FucT* o actividad de *XylT*, como se indica en otro punto de la presente memoria. De forma alternativa, los fragmentos de un polinucleótido que son útiles como sondas de hibridación generalmente no codifican proteínas fragmento que retengan actividad biológica. Los fragmentos de un polinucleótido de *FucT* o *XylT* también se pueden usar para diseñar secuencias inhibidoras para la supresión de la expresión del polipéptido de *FucT* y/o de *XylT*. Por tanto, por ejemplo, los fragmentos de una secuencia de nucleótidos pueden oscilar entre al menos aproximadamente 15 nucleótidos, 20 nucleótidos, aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos, aproximadamente 150 nucleótidos, aproximadamente 200 nucleótidos, aproximadamente 250 nucleótidos, aproximadamente 300 nucleótidos, aproximadamente 350 nucleótidos, aproximadamente 400 nucleótidos, aproximadamente 450 nucleótidos, aproximadamente 500 nucleótidos, aproximadamente 550 nucleótidos, aproximadamente 600 nucleótidos, aproximadamente 650 nucleótidos, aproximadamente 700 nucleótidos, aproximadamente 750 nucleótidos, aproximadamente 800 nucleótidos y hasta el polinucleótido de longitud completa que codifica las proteínas descritas en la presente memoria.

Un fragmento de un polinucleótido de *FucT* que codifica una porción biológicamente activa de una proteína de *FucT* de la invención codificará al menos 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 475, 500 aminoácidos

contiguos, o hasta el número total de aminoácidos presentes en una proteína de FucT de longitud completa de la invención (por ejemplo, 509 aminoácidos para la SEC ID N°: 3). Un fragmento de un polinucleótido de XylT que codifica una porción biológicamente activa de una proteína de XylT de la invención codificará al menos 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 475 aminoácidos contiguos, o hasta el número total de aminoácidos presentes en una proteína de XylT de longitud completa de la invención (por ejemplo, 490 aminoácidos para la SEC ID N°: 3). Un fragmento de un polinucleótido de XylT que codifica una porción biológicamente activa de una proteína de XylT de longitud parcial de la invención codificará al menos 15, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, aminoácidos contiguos, o hasta el número total de aminoácidos presentes en una proteína de XylT de longitud parcial de la invención (por ejemplo, 490 aminoácidos para la SEC ID N°: 21).

Por tanto, un fragmento de un polinucleótido de FucT o de XylT puede codificar una porción biológicamente activa de una proteína de FucT o XylT, respectivamente, o puede ser un fragmento que se puede usar como sonda de hibridación o como cebador de PCR, o se puede usar para diseñar secuencias inhibidoras para la supresión; usando los métodos descritos a continuación. Se puede preparar una porción biológicamente activa de una proteína de FucT o de XylT aislando una porción de uno de los polinucleótidos de FucT o XylT de la invención, respectivamente, expresando la porción codificada de la proteína de FucT o XylT (por ejemplo, mediante expresión recombinante *in vitro*) y determinando la actividad de la porción codificada del polipéptido de FucT o XylT. Los polinucleótidos que son fragmentos de una secuencia de nucleótidos de FucT o XylT comprenden al menos 15, 20, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 800, 900, 1.000, 1.100, 1.200, 1.300, 1.400 ó 1.450 nucleótidos contiguos, o hasta el número de nucleótidos presentes en un polinucleótido de FucT o XylT descrito en la presente memoria (por ejemplo, 1865, 1473, 1860, 1530, 1282 ó 1275 nucleótidos correspondientes a las SEC ID N°: 1, 2, 4, 5, 19 y 20, respectivamente).

“Variantes” pretende significar secuencias sustancialmente similares. Para polinucleótidos, una variante comprende una eliminación y/o adición de uno o más nucleótidos en uno o más sitios dentro del polinucleótido nativo y/o una sustitución de uno o más nucleótidos en uno o más sitios en el polinucleótido nativo. Tal como se usa en la presente memoria, un polinucleótido o polipéptido “nativo” comprende una secuencia de nucleótidos o una secuencia de aminoácidos natural, respectivamente. Para polinucleótidos, las variantes conservativas incluyen aquellas que, debido a la degeneración del código genético, codifican la secuencia de aminoácidos de uno de los polipéptidos de FucT o XylT de la invención. Las variantes alélicas naturales como las indicadas pueden identificarse con el uso de técnicas de biología molecular bien conocidas, como, por ejemplo, con técnicas de reacción en cadena de polimerasa (PCR) y de hibridación como se describe más adelante. Las variantes de polinucleótidos también incluyen polinucleótidos derivados sintéticamente, tal como los generados, por ejemplo, usando mutagénesis dirigida a sitio pero que todavía codifican una proteína FucT o XylT de la invención. Generalmente, las variantes de un polinucleótido particular (por ejemplo, la SEC ID N°: 1, la SEC ID N°: 2, la SEC ID N°: 4, la SEC ID N°: 5, la SEC ID N°: 19 ó la SEC ID N°: 20, los fragmentos de los mismos, y los complementos de dichas secuencias) tendrán al menos aproximadamente un 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% o más de identidad de secuencia con dicho polinucleótido particular, determinado mediante programas de alineamiento de secuencia y parámetros descritos en otro sitio de la presente memoria.

Las variantes de un polinucleótido particular (es decir, el polinucleótido de referencia) también puede evaluarse por comparación del porcentaje de identidad de secuencia entre el polipéptido codificado por un polinucleótido variante y el polipéptido codificado por el polinucleótido de referencia. Por tanto, por ejemplo, se describe un polinucleótido aislado que codifica un polipéptido con un porcentaje dado de identidad de secuencia con el polipéptido de FucT o XylT de SEC ID N°: 3, SEC ID N°: 6 ó SEC ID N°: 21, respectivamente. El porcentaje de identidad de secuencia entre dos polipéptidos cualquiera se puede calcular usando programas de alineamiento de secuencia y los parámetros descritos en otra parte de la presente memoria. Cuando se evalúa cualquier par dado de polinucleótidos por comparación del porcentaje de secuencia de identidad compartida por los dos polipéptidos que codifican, el porcentaje de identidad de secuencia entre dos polipéptidos codificados es al menos aproximadamente 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ó más de identidad de secuencia.

Proteína “variante” pretende significar una proteína derivada de la proteína nativa por eliminación o adición de uno o más aminoácidos en uno o más sitios de la proteína nativa y/o sustitución de uno o más aminoácidos en uno o más sitios de la proteína nativa. Las proteínas variantes son biológicamente activas, es decir, continúan poseyendo la actividad biológica deseada de la proteína nativa, es decir, la actividad enzimática de unir el residuo de fucosa α 1,3-ligada (actividad de FucT) o residuo de xilosa β 1,2-ligada (actividad de XylT) a N-glicanos de glicoproteína en plantas como se describe en la presente memoria. Dichas variantes pueden resultar de, por ejemplo, polimorfismo genético o a partir de manipulación humana. Las variantes biológicamente activas de una proteína FucT o XylT nativas tendrán al menos aproximadamente 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ó más de identidad de secuencia con la secuencia de aminoácidos correspondiente a la proteína nativa, determinado mediante programas de alineamiento de secuencia y parámetros descritos en otro sitio de la presente memoria. Una variante biológicamente activa de una proteína puede diferir de dicha proteína en solo 1-15 residuos de aminoácido, solo 1-10, tal como 6-10, solo 5, solo 4, 3, 2 ó incluso 1 residuo de aminoácido.

Las proteínas se pueden alterar de diferentes modos, que incluyen sustituciones, eliminaciones, truncados e

inserciones de aminoácidos. Los métodos para dichas manipulaciones se conocen de forma general en la técnica. Por ejemplo, las variantes y fragmentos de secuencias de aminoácidos de las proteínas de FucT y XylT se pueden preparar mediante mutaciones del ADN. Los métodos de mutagénesis y alteraciones de polinucleótidos son bien conocidos en la técnica. Véase, por ejemplo, Kunkel (1985) Proc. Natl. Acad. Sci. USA 82: 488-492; Kunkel et al. (1987) Methods in Enzymol. 154: 367-382; Patente de EE.UU. N° 4.873.192; Walker y Gaastra, eds. (1983) "Techniques in Molecular Biology" (MacMillan Publishing Company, Nueva York) y las referencias citadas en ellas. Una guía para realizar sustituciones de aminoácidos apropiadas que no afecten a la actividad biológica de la proteína de interés puede encontrarse en el modelo de Dayhoff et al. (1978) "Atlas of Protein Sequence and Structure" (Natl. Biomed. Res. Found., Washington, D.C.). Las sustituciones conservativas, tales como intercambiar un aminoácido por otro que tenga propiedades similares, pueden ser óptimas.

De este modo, los polinucleótidos incluyen tanto secuencias de FucT y XylT naturales como formas mutantes. Asimismo, las proteínas abarcan tanto proteínas de FucT y XylT naturales como variaciones y formas modificadas de las mismas. Dichas variantes continuarán poseyendo la actividad deseada. Por tanto, cuando se desea la expresión de una proteína funcional, la proteína expresada poseerá la actividad de FucT o XylT deseada, es decir, la actividad enzimática de unir el residuo de fucosa α 1,3-ligada (actividad de FucT) o el residuo de xilosa β 1,2-ligada (actividad de XylT) a N-glicanos de glicoproteína en plantas como se describe en la presente memoria. Cuando el objetivo es la inhibición de la expresión o la función del polipéptido de FucT y/o XylT, la actividad deseada del polinucleótido o polipéptido variante es una de inhibición de expresión o función del respectivo polipéptido de FucT y/o XylT. Obviamente, cuando se desea la expresión de una variante FucT o XylT funcional, las mutaciones que se harán en el ADN que codifica la variante no deben colocar la secuencia fuera del marco de lectura y de manera óptima no crearán regiones complementarias que podrían producir una estructura de ARNm secundaria. Véase, la Publicación de Solicitud de Patente EP N° 75.444.

Cuando se desea una proteína funcional, no se espera que las eliminaciones, inserciones y sustituciones de las secuencias de proteína abarcadas en la presente memoria produzcan cambios radicales en las características de la proteína. Sin embargo, cuando es difícil predecir el efecto exacto de la sustitución, eliminación o inserción antes de realizarla, el especialista en la técnica apreciará que el efecto se evaluará mediante ensayos de escrutinio rutinarios, que incluyen los ensayos para monitorizar la actividad de FucT y XylT descritos en la presente memoria más adelante en la sección Experimental.

Las polinucleótidos y proteínas variantes también abarcan secuencias y proteínas derivadas de un procedimiento mutagénico y recombinogénico tal como el barajado de ADN. Con dicho procedimiento, se pueden manipular una o más secuencias codificadoras de FucT o XylT diferentes para crear una nueva proteína de FucT o de XylT que posea las propiedades deseadas. De esta manera, se generan bibliotecas de polinucleótidos recombinantes a partir de una población de polinucleótidos de secuencia relacionada que comprenden regiones de secuencia que tienen una identidad de secuencia sustancial y que pueden ser recombinadas homogéneamente *in vitro* o *in vivo*. Por ejemplo, usando esta estrategia, las estructuras de secuencia que codifican un dominio de interés se pueden barajar entre el gen de FucT o XylT de la invención y otros genes de FucT o XylT conocidos, respectivamente, para obtener un nuevo gen que codifica para una proteína con una propiedad de interés mejorada. Las estrategias para dicho barajado de ADN son conocidas en la técnica. Véase, por ejemplo, Stemmer (1994) Proc. Natl. Acad. Sci. USA 91: 10747-10751; Stemmer (1994) Nature 370: 389-391; Crameri et al. (1997) Nature Biotech. 15: 436-438; Moore et al. (1997) J. Mol. Biol. 272: 336-347; Zhang et al. (1997) Proc. Natl. Acad. Sci. USA 94: 4504-4509; Crameri et al. (1998) Nature 391: 288-291; y las Patentes de EE.UU. N° 5.605.793 y 5.837.458.

La comparación de secuencias y la determinación del porcentaje de identidad y de la similitud de identidad entre dos secuencias se puede llevar a cabo usando un algoritmo matemático. En una realización preferida, se determina el porcentaje de identidad entre dos secuencias de aminoácidos usando el algoritmo de Needleman y Wunsch (1970) J. Mol. Biol. 48: 444-453, que se incorpora al programa GAP en el paquete de software GCG (disponible en www.accelrys.com), usando una matriz BLOSSUM62 o una matriz PAM250, y un peso de hueco de 16, 14, 12, 10, 8, 6 ó 4, y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina usando el programa GAP del paquete de software GCG, usando una matriz de puntuación BLOSUM62 (véase Henikoff et al. (1989) Proc. Natl. Acad. Sci. USA 89: 10915) y un peso de hueco de 40, 50, 60, 70 u 80, y un peso de longitud de 1, 2, 3, 4, 5 ó 6. Una configuración de parámetros particularmente preferida (y la que debería usarse si el usuario no tiene claro qué parámetros deberían aplicarse para determinar si una molécula está dentro del límite de identidad de secuencia de la invención) es usando una matriz de puntuación BLOSUM62 con un peso de hueco de 60 y un peso de longitud de 3).

El porcentaje de identidad entre dos secuencias de aminoácidos o de nucleótidos también puede determinarse usando el algoritmo de Meyers y Miller (1989) CABIOS 4: 11-17, que se ha incorporado al programa ALIGN (versión 2.0), usando una tabla de residuos de peso PAM120, una penalización de longitud de hueco de 12 y una penalización de hueco de 4.

Una indicación alternativa de que dos moléculas de ácido nucleico están relacionadas estrechamente es que las dos moléculas se hibriden entre ellas en condiciones de severidad. Las condiciones de severidad dependen de las secuencias y son diferentes bajo diferentes parámetros ambientales. Generalmente, se seleccionan las condiciones de severidad para que sean aproximadamente de 5°C a 20°C menores que el punto de fusión térmico (T_m)

correspondiente a la secuencia específica a una fuerza iónica y un pH definidos. La T_m es la temperatura (a una fuerza iónica y un pH definidos) a la cual el 50% de la secuencia diana se hibrida con una sonda perfectamente coincidente. Las condiciones para la hibridación de ácido nucleico y para el cálculo de severidad se pueden encontrar, por ejemplo, en Sambrook et al. (2001) "Molecular Cloning: A Laboratory Manual" (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY) y en Tijssen (1993) "Hybridization With Nucleic Acid Probes, Part I: Theory and Nucleic Acid Preparation" (Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Ltd., NY, NY).

Para los propósitos de la presente invención, "condiciones severas" abarcan las condiciones en las que la hibridación se producirá solo si hay menos de un 25% de discordancia entre la molécula de hibridación y la secuencia diana. Las "condiciones severas" se pueden dividir en niveles concretos de severidad para una definición más precisa. Por tanto, tal como se usan en la presente memoria, las condiciones de "severidad moderada" son aquellas en las que las moléculas con más de un 25% de discordancia no se hibridarán; las condiciones de "severidad media" son aquellas en las que las moléculas con más del 15% de discordancia no se hibridarán, y las condiciones de "alta severidad" son aquellas en las que las secuencias con más del 10% de discordancia no se hibridarán. Las condiciones de "severidad muy elevada" son aquellas en las que las secuencias con más del 6% de discordancia no se hibridarán.

Los polinucleótidos de FucT y XylT descritos en la presente memoria se pueden usar como sondas para el aislamiento de las correspondientes secuencias homólogas en otros organismos, más particularmente en otras especies de plantas. De esta manera, métodos tales como la PCR, la hibridación y otros similares se pueden usar para identificar secuencias basadas en su homología de secuencia con las secuencias descritas en la presente memoria. Véase, por ejemplo, Sambrook et al. (1989) "Molecular Cloning: A Laboratory Manual" (2ª edición, Cold Spring Harbor Laboratory Press, Plainview, Nueva York) e Innis et al. (1990), "PCR Protocols: A Guide to Methods and Applications" (Academic Press, Nueva York). Por tanto, se pueden aislar secuencias de polinucleótidos en base a su identidad de secuencia con los polinucleótidos completos de FucT o XylT de la presente memoria (es decir, las SEC ID N°: 1 y 2 para FucT; las SEC ID N°: 4 y 5 para la isoforma n°1 de XylT de SEC ID N°: 6, y las SEC ID N°: 19 y 20 para la isoforma n°2 de XylT de SEC ID N°: 21) o con los fragmentos y variantes de los mismos.

En un método de PCR, se pueden diseñar los cebadores de oligonucleótidos para su uso en reacciones PCR para la amplificación de las correspondientes secuencias de ADN a partir de ADNc o ADN genómico extraído de cualquier organismo de interés. Los métodos conocidos de PCR incluyen, aunque sin limitación, métodos que usan cebadores emparejados, cebadores anidados, cebadores específicos sencillos, cebadores degenerados, cebadores específicos de gen, cebadores específicos de vector, cebadores parcialmente discordantes, y otros similares. Los métodos para diseñar cebadores de PCR y para clonación PCR son conocidos de manera general en la técnica y se describen en Sambrook et al. (1989) "Molecular Cloning: A Laboratory Manual" (2ª edición, Cold Spring Harbor Laboratory Press, Plainview, Nueva York); Innis y Gelfand, editores (1995) "PCR Strategies" (Academic Press, Nueva York) e Innis y Gelfand, editores (1999) "PCR Methods Manual" (Academic Press, Nueva York).

En un método de hibridación, la totalidad o parte de una secuencia de nucleótidos conocida se puede usar como sonda que se hibrida selectivamente con otros polinucleótidos correspondientes presentes en una población de fragmentos de ADN genómico o fragmentos de ADNc (es decir, bibliotecas de ADNc o genómicas) de otro organismo de interés. Las denominadas sondas de hibridación pueden ser fragmentos de ADN genómico, fragmentos de ADNc, fragmentos de ARN u otros oligonucleótidos, y pueden marcarse con un grupo detectable tal como un ^{32}P , o cualquier otro marcador detectable. Las sondas para hibridación se pueden preparar marcando oligonucleótidos sintéticos en base a la secuencia de interés, por ejemplo, los polinucleótidos de FucT o XylT. Adicionalmente se pueden usar cebadores degenerados diseñados en base a nucleótidos o a residuos de aminoácidos conservados de la secuencia de nucleótidos o de aminoácidos codificada. Los métodos de construcción de bibliotecas de ADNc o genómicas, y de preparación de sondas de hibridación, son conocidos de forma general en la técnica y se describen en Sambrook et al. (1989) "Molecular Cloning: A Laboratory Manual" (2ª edición, Cold Spring Harbor Laboratory Press, Plainview, Nueva York).

Por ejemplo, la totalidad o parte de las secuencias de polinucleótidos de FucT o XylT específicas conocidas se puede usar como sonda que se hibride selectivamente con otros nucleótidos y ARNs mensajeros de FucT o XylT. Para alcanzar una hibridación específica en una variedad de condiciones, dichas sondas incluyen secuencias que son únicas y que preferiblemente tienen una longitud de al menos aproximadamente 10 nucleótidos, y de forma más óptima una longitud de al menos aproximadamente 20 nucleótidos. Esta técnica puede usarse para aislar otras secuencias de nucleótidos de FucT o XylT correspondientes a partir de un organismo deseado o como ensayo diagnóstico para determinar la presencia de una secuencia codificadora de FucT o XylT en un organismo. Las técnicas de hibridación incluyen el escrutinio de hibridación de bibliotecas de ADN en placa (en plaquetas o en colonias; véase, por ejemplo, Innis et al., editores (1990) "PCR Protocols: A Guide to Methods and Applications" (Academic Press, Nueva York).

Por lo tanto, además de los polinucleótidos de FucT y XylT nativos y de los fragmentos y variantes de los mismos, los polinucleótidos aislados en la presente memoria también abarcan secuencias de ADN homólogas identificadas y aisladas a partir de otros organismos mediante hibridación con secuencias enteras o parciales obtenidas a partir de polinucleótidos de FucT o XylT o de variantes de los mismos. Las condiciones que permitirán que otras secuencias

de ADN se hibriden con las secuencias de ADN descritas en la presente memoria se pueden determinar de acuerdo a técnicas conocidas de manera general en la técnica. Por ejemplo, la hibridación de dichas secuencias se puede llevar a cabo en diversas condiciones de severidad moderada, media, alta o muy alta, tal como se ha indicado anteriormente en la presente memoria.

5 Métodos de la Invención:

La presente invención está dirigida a métodos para alterar los modelos de glicosilación de proteínas en plantas de lentejas de agua, particularmente en plantas que actúan como hospedantes para la producción de proteínas recombinantes, particularmente proteínas recombinantes de mamífero de interés farmacéutico. Los métodos tienen utilidad para la producción de plantas de lentejas de agua que son capaces de producir proteínas recombinantes que tienen un modelo de *N*-glicosilación que se asemeja estrechamente al que se da en mamíferos. Las composiciones de la invención incluyen plantas de lentejas de agua que se transforman de manera estable para comprender un modelo de *N*-glicosilación alterado de sus proteínas endógenas (es decir, homólogas) y heterólogas producidas recombinantemente. En algunas realizaciones, las plantas son plantas de lentejas de agua transgénicas que producen anticuerpos monoclonales (mAbs) para proteínas de mamífero que tienen una actividad ADCC potenciada respecto a los mAbs producidos en una planta de control que no tiene la maquinaria de glicosilación alterada para reducir la unión específica de planta de residuos de α 1,3-fucosa a *N*-glicanos de las glicoproteínas homólogas y heterólogas producidas según la presente memoria.

Los métodos de la invención se dirigen a la supresión (es decir, la inhibición) de la expresión de una o ambas de las enzimas implicadas en la producción de glicoproteínas complejas de plantas superiores. De especial interés es la supresión de una fucosiltransferasa o una o más de sus isoformas, la supresión de una xilosiltransferasa o una o más isoformas de la misma, o la supresión de la expresión de ambas de estas proteínas y una o más isoformas de las mismas. Se reconoce que la supresión de la fucosiltransferasa y/o la xilosiltransferasa y una o más isoformas de las mismas se puede llevar a cabo de manera temporal. Alternativamente, suprimiendo de manera estable la expresión de la fucosiltransferasa y/o de la xilosiltransferasa, es posible producir plantas superiores transgénicas que pasan de generación en generación, tanto asexualmente como sexualmente, la capacidad de producir glicoproteínas que tienen un modelo de *N*-glicosilación que se asemeja muy estrechamente al que se da en mamíferos, más particularmente, en humanos. Esto proporciona de forma ventajosa la producción de glicoproteínas de mamífero recombinantes que tiene una unión reducida del residuo de xilosa β 1,2-ligado y/o del residuo de fucosa α 1,3-ligado a *N*-glicanos de la glicoproteína.

La inhibición de la expresión de una de estas proteínas, o de forma ventajosa de ambas, en una planta de lenteja de agua, se puede llevar a cabo usando cualquier método conocido en la técnica. De esta manera, se introduce en la célula hospedante de interés un polinucleótido que comprende una secuencia inhibidora para FucT, XylIT, o una combinación de las mismas. Para una supresión temporal, la secuencia inhibidora de FucT o XylIT puede ser ARN interferente pequeño (ARNsi) o micro ARN (miARN) sintetizado químicamente o producido *in vitro* que, cuando se introdujeran en la célula hospedante, inhibirían directamente, aunque transitoriamente, la FucT, la XylIT o una combinación de ambas, silenciando la expresión de dicho(s) producto(s) génico(s) diana(s).

Alternativamente, la supresión estable de la expresión de FucT, XylIT o de una combinación de las mismas es deseable como se ha indicado anteriormente en esta memoria. Así, en algunas realizaciones, la actividad del polipéptido de FucT o XylIT de la invención se reduce o se elimina transformando una célula vegetal con un casete de expresión que expresa un polinucleótido que inhibe la expresión de la FucT o XylIT, o ambas. Los métodos para inhibir o eliminar la expresión de un gen en una planta son bien conocidos en la técnica, y cualquier método de ese tipo se puede usar en la presente invención para inhibir la expresión de FucT o XylIT, o ambas.

Por tanto, en algunas realizaciones, la expresión de la proteína de FucT y/o XylIT se puede inhibir introduciendo en la planta una construcción de nucleótido, tal como una casete de expresión, que comprende una secuencia que codifica una molécula de nucleótido inhibidora que está diseñada para silenciar la expresión del producto génico de FucT y/o XylIT de interés, tal como ARN con orientación sentido, ARN antisentido, ARN de cadena doble (ARNds), ARN de horquilla (ARNhp), ARNhp que contiene intrón, ARN catalítico, miARN y otros similares. En algunas realizaciones, la construcción de nucleótidos, por ejemplo, una casete de expresión, puede comprender una secuencia que codifica un ARNm, cuya traducción da lugar a un polipéptido de interés que inhibe la expresión o la función del producto génico de FucT y/o XylIT de interés. Cuando la construcción de nucleótido comprende una secuencia que codifica una molécula de nucleótido inhibidora o un ARNm para un polipéptido de interés, la secuencia está ligada operativamente a un promotor que conduce la expresión en una célula vegetal de tal modo que la molécula de nucleótido inhibidora codificada o el ARNm pueden ser expresados.

De acuerdo con la presente invención, la expresión de un gen de FucT o XylIT se inhibe si el nivel de proteína de la FucT o la XylIT es estadísticamente inferior al nivel de proteína de la misma FucT o XylIT en una planta que no ha sido modificada genéticamente o mutagenizada para inhibir la expresión de dicha FucT o XylIT. En realizaciones particulares de la invención, el nivel de proteína de la FucT o la XylIT, o de ambas, en una planta modificada de acuerdo con la invención es inferior al 95%, inferior al 90%, inferior al 80%, inferior al 70%, inferior al 60%, inferior al 50%, inferior al 40%, inferior al 30%, inferior al 20%, inferior al 10%, o inferior al 5% del nivel de proteína de la misma FucT o XylIT en una planta que no es un mutante o que no ha sido modificada genéticamente para inhibir la

expresión de dicha FucT o XylT, o ambas, FucT y XylT. El nivel de expresión de la FucT o la XylT, o de ambas, se puede medir directamente, por ejemplo, realizando un ensayo de determinación del nivel de FucT o XylT, o ambas, expresadas en la célula vegetal o en la planta, o indirectamente, por ejemplo, observando el efecto en una planta transgénica a nivel fenotípico, es decir, mediante análisis de planta transgénica, observado como una reducción, o incluso la eliminación, de la unión de los residuos de β 1,2-xilosa y/o de α 1,3-fucosa con los *N*-glicanos de glicoproteínas de la planta.

En otras realizaciones de la invención, la actividad de FucT o XylT, o de ambas, se reduce o se elimina transformando una célula vegetal con una casete de expresión que comprende un polinucleótido que codifica un polipéptido que inhibe la actividad de FucT o XylT, o de ambas. La actividad de una FucT o XylT se inhibe de acuerdo a la presente invención si la actividad de la FucT o XylT es estadísticamente inferior a la actividad de la misma FucT o XylT en una planta que no ha sido modificada genéticamente para inhibir la actividad de dicha FucT o XylT. En realizaciones particulares de la invención, la actividad de la FucT o de la XylT en una planta modificada de acuerdo con la invención es inferior al 95%, inferior al 90%, inferior al 80%, inferior al 70%, inferior al 60%, inferior al 50%, inferior al 40%, inferior al 30%, inferior al 20%, inferior al 10% o inferior al 5% de la actividad de la misma FucT o XylT en una planta que no ha sido modificada genéticamente para inhibir la expresión de dicha FucT o XylT. La actividad de una FucT o XylT es "eliminada" de acuerdo a la invención cuando no es detectable mediante los métodos de ensayo descritos en otra parte de la presente memoria.

Por ejemplo, en algunas realizaciones, la secuencia inhibidora de FucT o la secuencia inhibidora de XylT (o ambas) se expresan en la orientación sentido, en donde los transcritos con orientación sentido causan la co-supresión de la expresión de una o ambas enzimas. Alternativamente, la secuencia inhibidora de FucT y/o XylT (por ejemplo, una secuencia de longitud completa, una secuencia truncada, fragmentos de una secuencia, combinaciones de los mismos, y otros similares) se pueden expresar en la orientación antisentido e inhibir de este modo la expresión o la función endógenas de FucT y/o XylT mediante mecanismos antisentido.

En otras realizaciones adicionales, la secuencia o secuencias inhibidoras de FucT y/o XylT se expresan como un ARN de horquilla, que comprende tanto una secuencia sentido como una secuencia antisentido. En las realizaciones que comprenden una estructura de horquilla, la estructura de lazo puede comprender cualquier secuencia de nucleótidos adecuada que incluya por ejemplo regiones 5' no traducidas y/o traducidas del gen que va a suprimirse, tal como 5' UTR y/o región traducida del polinucleótido de FucT de SEC ID N°: 1 ó 2, o la 5' UTR y/o región traducida del polinucleótido de XylT de SEC ID N°: 4, 5, 19 ó 20, y otras similares. En algunas realizaciones, la secuencia inhibidora de FucT o XylT expresada como una horquilla es codificada mediante una región invertida de la secuencia de nucleótidos de FucT o XylT. En otras realizaciones adicionales, las secuencias inhibidoras de FucT y/o XylT se expresan como ARN de cadena doble, en donde una secuencia inhibidora de FucT y/o XylT es expresada en la orientación sentido y otra secuencia complementaria es expresada en la orientación antisentido. El ARN de cadena doble, las estructuras de horquilla, y las combinaciones de las mismas que comprenden secuencias de nucleótidos de FucT, secuencias de nucleótidos de XylT, o combinaciones de las mismas, pueden operar mediante interferencia de ARN, co-supresión, mecanismo antisentido, cualquier combinación de las mismas, o a través de cualquier otro mecanismo que provoque la inhibición de la expresión o la función de FucT y/o XylT.

Por tanto, se pueden usar muchos métodos para reducir o eliminar la actividad de una FucT o XylT, o de ambas proteínas, y de cualquiera de sus isoformas. Mediante "isoforma" se pretende indicar una proteína variante natural de la proteína de FucT o XylT de interés, en donde la variante es codificada por un gen diferente. Generalmente, las isoformas de una proteína de FucT o XylT particular son codificadas mediante una secuencia de nucleótidos que tiene al menos un 90% de identidad de secuencia con respecto a la secuencia de nucleótidos que codifica para la proteína FucT o XylT de interés. Se puede usar más de un método para reducir o eliminar la actividad de una FucT o XylT vegetal individual, y de sus isoformas. A continuación se muestran ejemplos no limitantes de métodos de reducción o eliminación de la actividad de una FucT o XylT vegetal.

Métodos basados en Polinucleótidos:

En algunas realizaciones de la presente invención, se transforma una célula vegetal con una casete de expresión que es capaz de expresar un polinucleótido que inhibe la expresión de FucT o XylT, o de ambas. El término "expresión" tal como se usa en la presente memoria se refiere a la biosíntesis de un producto génico, que incluye la transcripción y/o la traducción de un producto génico. Por ejemplo, para los propósitos de la presente invención, una casete de expresión capaz de expresar un polinucleótido que inhibe la expresión de al menos una de FucT o XylT, o ambas, es una casete de expresión capaz de producir una molécula de ARN que inhibe la transcripción y/o la traducción de al menos una de FucT o XylT, o ambas. La "expresión" o la "producción" de una proteína o polipéptido a partir de una molécula de ADN se refiere a la transcripción y a la traducción de la secuencia codificadora para producir la proteína o polipéptido, mientras que la "expresión" o "producción" de una proteína o polipéptido a partir de una molécula de ARN se refiere a la traducción de la secuencia codificadora de ARN para producir la proteína o polipéptido.

A continuación se muestran ejemplos de polinucleótidos que inhiben la expresión de una FucT o XylT, o ambas.

Supresión/Co-supresión sentido:

En algunas realizaciones de la invención, la inhibición de la expresión de FucT o XylT, o de ambas, se puede obtener mediante supresión o co-supresión sentido. Para la co-supresión, se diseña una casete de expresión para expresar una molécula de ARN correspondiente a parte o a todo el ARN mensajero que codifica una FucT o XylT, o ambas, en la orientación "sentido". La sobreexpresión de la molécula de ARN puede dar como resultado una expresión reducida del gen nativo. Por consiguiente, se escrutan múltiples líneas vegetales transformadas con la casete de expresión de co-supresión para identificar aquellas que muestran la mayor inhibición de la expresión de FucT o XylT.

El polinucleótido usado para co-supresión puede responder a parte o a toda la secuencia que codifica la FucT o la XylT, parte o toda la región no traducida 5' y/o 3' de un transcrito de FucT o XylT, parte o toda la secuencia codificadora y de las regiones no traducidas de un transcrito que codifica FucT o XylT. En algunas realizaciones en las que el polinucleótido comprende parte o toda la región codificadora para la proteína de FucT o XylT, la casete de expresión se diseña para eliminar el codón de inicio del polinucleótido de tal modo que no transcribirá ningún producto de proteína.

Se puede usar la co-supresión para inhibir la expresión de genes vegetales para producir plantas que tengan niveles de proteínas indetectables para las proteínas codificadas por dichos genes. Véase, por ejemplo, Broin et al. (2002) *Plant Cell* 14: 1417-1432. La co-supresión también se puede usar para inhibir la expresión de proteínas múltiples en la misma planta. Véase, por ejemplo, la Patente de EE.UU. N° 5.942.657. Los métodos para usar la co-supresión para inhibir la expresión de genes endógenos en plantas se describen en Flavell et al. (1994) *Proc. Natl. Acad. Sci. USA* 91: 3490-3496; Jorgensen et al. (1996) *Plant Mol. Biol.* 31: 957-973; Johansen y Carrington (2001) *Plant Physiol.* 129: 1723-1731; Yu et al. (2003) *Phytochemistry* 63: 753-763; y las Patentes de EE.UU. N° 5.034.323, 5.283.184 y 5.942.657. La eficacia de la co-supresión puede aumentarse incluyendo una región poli-dT en la casete de expresión en una posición 3' respecto a la secuencia sentido y 5' respecto a la señal de poliadenilación. Véase la Publicación de Patente de EE.UU. N° 20020048814. Típicamente, dicha secuencia de nucleótidos tiene una identidad de secuencia sustancial con respecto a la secuencia del transcrito del gen endógeno, de manera óptima superior a aproximadamente un 65% de identidad de secuencia, más óptimamente superior a aproximadamente un 85% de identidad de secuencia, lo más óptimamente superior a aproximadamente un 95% de identidad de secuencia. Véase, por ejemplo, las Patentes de EE.UU. N° 5.283.184 y 5.034.323.

Supresión Antisentido:

En algunas realizaciones de la invención, la inhibición de la expresión de FucT o XylT, o de ambas, se puede obtener mediante supresión antisentido. Para la supresión antisentido, la casete de expresión se diseña para expresar una molécula de ARN complementaria a parte o a todo el ARN mensajero que codifica el FucT o XylT. La sobreexpresión de la molécula de ARN antisentido puede dar como resultado una expresión reducida del gen nativo. Por consiguiente, se escrutan múltiples líneas vegetales transformadas con la casete de expresión de supresión antisentido para identificar aquellas que muestren la mayor inhibición de expresión de FucT o XylT.

El polinucleótido para uso en supresión antisentido puede corresponder a parte o a todo el complemento de la secuencia que codifica la FucT o XylT, parte o todo el complemento de la región no traducida 5' y/o 3' del transcrito de FucT o XylT, o parte o todo el complemento de la secuencia codificadora y las regiones no traducidas de un transcrito que codifica la FucT o la XylT. Adicionalmente, el polinucleótido antisentido puede ser completamente complementario (es decir, 100% idéntico al complemento de la secuencia diana) o parcialmente complementario (es decir, menos de un 100% idéntico al complemento de la secuencia diana) a la secuencia diana. Se puede usar la supresión antisentido para inhibir la expresión de proteínas múltiples en la misma planta. Véase, por ejemplo, la Patente de EE.UU. N° 5.942.657. Además, se pueden usar porciones de los nucleótidos antisentido para romper la expresión del gen diana. Generalmente, se pueden usar secuencias de al menos 50 nucleótidos, 100 nucleótidos, 200 nucleótidos, 300, 400, 450, 500, 550 o más. Los métodos para usar la supresión antisentido para inhibir la expresión de genes endógenos en plantas se describen, por ejemplo, en Liu et al. (2002) *Plant Physiol.* 129: 1732-1743 y las Patentes de EE.UU. N° 5.759.829 y 5.942.657. La eficiencia de la supresión antisentido puede aumentarse incluyendo una región de poli-dT en la casete de expresión en la posición 3' respecto a la secuencia antisentido y 5' respecto a la señal de poliadenilación. Véase la Publicación de Patente a EE.UU. N° 20020048814.

Interferencia de ARN de Cadena Doble:

En algunas realizaciones de la invención, la inhibición de la expresión de una FucT o XylT, o de ambas, se puede obtener mediante interferencia de ARN de cadena doble (ARNds). Para la interferencia de ARNds, una molécula de ARN sentido como la descrita anteriormente para co-supresión y una molécula de ARN antisentido que es total o parcialmente complementaria a la molécula de ARN sentido se expresan en la misma célula, dando como resultado la inhibición de la expresión del ARN mensajero endógeno correspondiente.

La expresión de las moléculas sentido y antisentido puede lograrse diseñando la casete de expresión para que comprenda una secuencia sentido y una secuencia antisentido. Alternativamente, se pueden usar casetes de expresión separadas para las secuencias sentido y antisentido. A continuación se escrutan múltiples líneas vegetales transformadas con la casete de expresión o casetes de expresión de interferencia de ARNds para identificar líneas vegetales que muestren la mayor inhibición de la expresión de FucT o de XylT. Los métodos para

usar interferencia de ARNs para inhibir la expresión de genes de plantas endógenos se describen en Waterhouse et al. (1998) Proc. Natl. Acad. Sci. USA 95: 13959-13964, Liu et al. (2002) Plant Physiol. 129: 1732-1743 y los documentos WO 99/49029, WO 99/53050, WO 99/61631 y WO 00/49035.

Interferencia de ARN de horquilla e Interferencia de ARN de horquilla que contiene intrón

- 5 En algunas realizaciones de la invención, la inhibición de la expresión de FucT o XylIT, o de ambas, se puede obtener mediante interferencia de ARN de horquilla (ARNhp) o mediante interferencia de ARN de horquilla que contiene intrón (ARNihp). Estos métodos son altamente efectivos para inhibir la expresión de genes endógenos. Véase, Waterhouse y Helliwell (2003) Nat. Rev. Genet. 4: 29-38, y las referencias citadas en dicho documento.

10 Para la interferencia de ARNhp, se diseña la casete de expresión para expresar una molécula de ARN que se hibride consigo misma para formar una estructura de horquilla que comprenda una región de lazo de cadena sencilla y un tallo de bases emparejadas. La región de tallo de bases emparejadas comprende una secuencia sentido correspondiente a parte o a todo el ARN mensajero endógeno que codifica el gen cuya expresión se va a inhibir, y una secuencia antisentido que es total o parcialmente complementaria a la secuencia sentido. Así, la región de tallo de bases emparejadas de la molécula generalmente determina la especificidad de la interferencia de ARN. Las moléculas de ARNhp son altamente efectivas en la inhibición de la expresión de genes endógenos, y la interferencia de ARN que inducen es heredada por las generaciones de plantas posteriores. Véase, por ejemplo, Chuang y Meyerowitz (2000) Proc. Natl. Acad. Sci. USA 97: 4985-4990; Stoutjesdijk et al. (2002) Plant Physiol. 129: 1723-1731; y Waterhouse y Helliwell (2003) Nat. Rev. Genet. 4: 29-38. Los métodos para usar la interferencia de ARNhp para inhibir o silenciar la expresión de genes se describen, por ejemplo, en Chuang y Meyerowitz (2000) Proc. Natl. Acad. Sci. USA 97: 4985-4990; Stoutjesdijk et al. (2002) Plant Physiol. 129: 1723-1731; Waterhouse y Helliwell (2003) Nat. Rev. Genet. 4: 29-38; Pandolfini et al. BMC Biotechnology 3: 7 y la Publicación de Patente de EE.UU. N° 20030175965. Un ensayo temporal para determinar la eficacia de construcciones de ARNhp para silenciar la expresión génica in vivo se describe en Panstruga et al. (2003) Mol. Biol. Rep. 30: 135-140.

25 Para ARNihp, las moléculas interferentes tiene la misma estructura general que para ARNhp, pero la molécula de ARN comprende además un intrón que es capaz de dividirse en la célula en la que se expresa el ARNihp. El uso de un intrón minimiza el tamaño del lazo de la molécula de ARN de horquilla después de la división, y esto aumenta la eficacia de la interferencia. Véase, por ejemplo, Smith et al. (2000) Nature 407: 319-320. De hecho, Smith et al. mostraron una supresión del 100% de la expresión génica endógena usando interferencia mediada por ARNihp. Los métodos para usar interferencia de ARNihp para inhibir la expresión de genes vegetales endógenos se describen, por ejemplo, en Smith et al. (2000) Nature 407: 319-320; Wesley et al. (2001) Plant J. 27: 581-590; Wang y Waterhouse (2001) Curr. Opin. Plant Biol. 5: 146-150; Waterhouse y Helliwell (2003) Nat. Rev. Genet. 4: 29-38; Helliwell y Waterhouse (2003) Methods 30: 289-295, y en la Publicación de Patente de EE.UU. N° 20030180945.

35 La casete de expresión correspondiente a interferencia de ARNhp también se puede diseñar de tal modo que la secuencia sentido y la secuencia antisentido no correspondan a un ARN endógeno. En esta realización, las secuencias sentido y antisentido flanquean a una secuencia de lazo que comprende una secuencia de nucleótido que corresponde a parte o a todo el ARN mensajero endógeno del gen diana. De este modo, es la región de lazo la que determina la especificidad de la interferencia de ARN. Véase, por ejemplo, el documento WO 02/00904.

40 El silenciamiento transcripcional de gen (TGS) se puede llevar a cabo a través de uso de construcciones de ARNhp en las que la repetición invertida de la horquilla comparte identidad de secuencia con la región promotora de un gen que va a ser silenciado. El procesamiento del ARNhp en ARNs cortos que pueden interactuar con la región promotora homóloga puede activar la degradación o la metilación para dar como resultado el silenciamiento (Aufsatz et al. (2002) PNAS 99 (Supl. 4): 16499-16506; Mette et al. (2000) EMBO J 19(19): 5194-5201).

45 Las casetes de expresión que se diseñan para expresar una molécula de ARN que forma una estructura de horquilla se denominan en la presente memoria casetes de expresión de ARNi. En algunas realizaciones, la casete de expresión de ARNi se diseña de acuerdo a una estrategia esbozada en la Figura 28. En dichas realizaciones, se puede diseñar una casete de expresión de ARNi para suprimir la expresión de los genes de FucT y XylIT individuales (es decir, cada casete proporciona un bloque de gen individual), o puede diseñarse para suprimir la expresión de ambos genes, de FucT y XylIT (es decir, una casete de expresión de ARNi individual expresa una molécula inhibidora que proporciona la supresión de la expresión de ambos genes). Cuando la casete de expresión de ARNi suprime la expresión de ambos genes, de FucT y de XylIT, se denomina en la presente memoria casete de expresión de ARNi "quimérica". Las casetes de expresión de ARNi de gen individual y quiméricas se pueden diseñar para expresar estructuras de ARNhp de mayor tamaño o, alternativamente, estructuras de ARNhp pequeñas, tal como se indica más adelante en la presente memoria.

55 Por tanto, en algunas realizaciones, la casete de expresión de ARNi se diseña para expresar estructuras de ARNhp más grandes que tengan una homología suficiente con el transcrito de ARNm diana para proporcionar el silenciamiento post-traducciona del gen de una o ambas de FucT y XylIT. Para estructuras de ARNhp más grandes, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés, un fragmento directo de secuencia génica de FucT o XylIT que comprende aproximadamente 500 a aproximadamente 800 nucleótidos (nt) de una cadena sentido

correspondiente a FucT o XylT, respectivamente, una secuencia espaciadora que comprende aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia como se indica más adelante en la presente memoria, y un fragmento inverso de la secuencia génica de XylT o FucT, en donde el fragmento inverso comprende la secuencia antisentido complementaria al respectivo fragmento directo (es decir, FucT o XylT). De este modo, por ejemplo, si un fragmento directo se representa mediante los nucleótidos "...acttg...", el correspondiente fragmento inverso se representa mediante los nucleótidos "...caagt...", y la cadena sentido correspondiente a dicha casete de expresión de ARNi comprendería la siguiente secuencia: "5'...acttg...nnnn...caagt...3'", en donde "nnnn" representa la secuencia espaciadora.

Se reconoce que el fragmento directo puede comprender una secuencia de nucleótidos que es un 100% idéntica a la porción correspondiente a la cadena sentido de la secuencia génica diana de FucT o XylT, o alternativamente, puede comprender una secuencia que comparte una identidad de al menos el 90%, al menos el 95% o al menos el 98% con respecto a la porción correspondiente de la cadena sentido del gen de FucT o XylT que va a ser silenciado. De forma similar, se reconoce que el fragmento inverso no comparte una identidad de secuencia del 100% con respecto al complemento del fragmento directo; si no que más bien debe tener una longitud suficiente y una complementariedad suficiente con la secuencia del fragmento directo, de tal modo que cuando la molécula de ARN inhibidora se exprese, las regiones transcritas correspondientes al fragmento directo y al fragmento inverso se hibridarán para formar un tallo de bases emparejadas (es decir, una porción de cadena doble) de la estructura de ARN hp. Por "longitud suficiente" se pretende indicar una longitud que es al menos un 10%, al menos un 15%, al menos un 20%, al menos un 30%, al menos un 40% de la longitud del fragmento directo, más frecuentemente al menos un 50%, al menos un 75%, al menos un 90% o al menos un 95% de la longitud del fragmento directo. Por "complementariedad suficiente" se pretende indicar la secuencia del fragmento comparte al menos un 90%, al menos un 95%, al menos un 98% de identidad de secuencia con el complemento de dicha porción del fragmento directo que se hibridará con el fragmento inverso para formar el tallo de bases emparejadas de la estructura de ARN hp. Por tanto, en algunas realizaciones, el fragmento inverso es el complemento (es decir, la versión antisentido) del fragmento directo.

Al diseñar dicha casete de expresión de ARNi, las longitudes del fragmento directo, la secuencia espaciadora y los fragmentos inversos se eligen de tal modo que la longitud combinada del polinucleótido que codifica la construcción de ARNhp es de aproximadamente 650 a aproximadamente 2500 nt, de aproximadamente 750 a aproximadamente 2500 nt, de aproximadamente 750 a aproximadamente 2400 nt, de aproximadamente 1000 a aproximadamente 2400 nt, de aproximadamente 1200 a aproximadamente 2300 nt, de aproximadamente 1250 a aproximadamente 2100 nt o de aproximadamente 1500 a aproximadamente 1800. En algunas realizaciones, la longitud combinada de la construcción de horquilla expresada es de aproximadamente 650 nt, aproximadamente 700 nt, aproximadamente 750 nt, aproximadamente 800 nt, aproximadamente 850 nt, aproximadamente 900 nt, aproximadamente 950 nt, aproximadamente 1000 nt, aproximadamente 1050 nt, aproximadamente 1100 nt, aproximadamente 1150 nt, aproximadamente 1200 nt, aproximadamente 1250 nt, aproximadamente 1300 nt, aproximadamente 1350 nt, aproximadamente 1400 nt, aproximadamente 1450 nt, aproximadamente 1500 nt, aproximadamente 1550 nt, aproximadamente 1600 nt, aproximadamente 1650 nt, aproximadamente 1700 nt, aproximadamente 1750 nt, aproximadamente 1800 nt, aproximadamente 1850 nt, aproximadamente 1900 nt, aproximadamente 1950 nt, aproximadamente 2000 nt, aproximadamente 2050 nt, aproximadamente 2100 nt, aproximadamente 2150 nt, aproximadamente 2200 nt, aproximadamente 2250 nt, aproximadamente 2300 nt, o cualquier longitud entre aproximadamente 650 nt y aproximadamente 2300 nt.

En algunas realizaciones, el fragmento directo comprende aproximadamente 500 a aproximadamente 800 nt, por ejemplo, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775 ó 800 nt de una cadena sentido correspondiente a FucT o XylT, por ejemplo, de la cadena sentido establecida en la SEC ID N°: 1 ó 2 (FucT) o las SEC ID N°: 4, 5, 19 ó 20 (XylT); la secuencia espaciadora comprende de aproximadamente 100 a aproximadamente 700 nt, por ejemplo, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675 ó 700 nt de cualquier secuencia como se indica más adelante, y el fragmento inverso comprende la cadena antisentido correspondiente a la secuencia de fragmento directo, o una secuencia que tiene una longitud suficiente y una complementariedad suficiente con la secuencia de fragmento directo.

La secuencia espaciadora puede ser cualquier secuencia que tenga una homología insuficiente con el gen diana, es decir, FucT o XylT, y una homología insuficiente consigo misma de tal modo que la porción de la molécula de ARN inhibidora expresada correspondiente a la región espaciadora no pueda auto-hibridarse, y por tanto forme el lazo de la estructura de ARN de horquilla. En algunas realizaciones, la secuencia espaciadora comprende un intrón, y por tanto la molécula de ARN inhibidora expresada forma un ARNihp tal como se ha indicado anteriormente en la presente memoria. En otras realizaciones, la secuencia espaciadora comprende una porción de la cadena sentido correspondiente al gen de FucT o XylT que va a ser silenciado, por ejemplo, una porción de la cadena sentido establecida en la SEC ID N°: 1 ó 2 (FucT) o las SEC ID N°: 4, 5, 19 ó 20 (XylT), particularmente una porción de la cadena sentido inmediatamente por debajo (dirección 3') de la secuencia de fragmento directo.

El promotor ligado operativamente puede ser cualquier promotor de interés que proporciona la expresión del polinucleótido inhibidor ligado operativamente dentro de la planta de interés, incluyendo uno de los promotores descritos más adelante en la presente memoria. La región reguladora puede comprender elementos regulatorios adicionales que potencian la expresión del polinucleótido inhibidor, que incluyen, aunque sin limitación, las

secuencias líder 5' y las secuencias líder 5' más intrones vegetales discutidos más adelante en la presente memoria.

En una realización, la casete de expresión de ARNi se diseña para suprimir la expresión del polipéptido de FucT de SEC ID N°: 3, una variante biológicamente activa del polipéptido de FucT de SEC ID N°: 3, o un polipéptido de FucT codificado por una secuencia que tiene al menos una identidad de secuencia del 75% con la secuencia de la SEC ID N°: 1 o la SEC ID N°: 2, por ejemplo, una identidad de secuencia de al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90% o al menos un 95% con respecto a la secuencia de la SEC ID N°: 1 o la SEC ID N°: 2. De esta manera, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo de la secuencia génica de FucT, en donde el fragmento directo comprende los nt 255-985 de la SEC ID N°: 1, una secuencia espaciadora que comprende aproximadamente de 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso de la secuencia génica de FucT, en donde el fragmento inverso comprende el complemento (es decir, versión antisentido) de los nt 255-985 de la SEC ID N°: 1. En dicha realización, la secuencia espaciadora se representa mediante los nt 986-1444 de la SEC ID N°: 1, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi correspondiente a la secuencia codificadora para la estructura de ARNhp es de 1918 nt. Transformar establemente una planta con una construcción de nucleótidos que comprende esta casete de expresión de ARNi, por ejemplo, el vector mostrado en la Figura 8, inhibe de manera efectiva la expresión de la FucT dentro de las células vegetales de la planta en la que la estructura de ARNhp es expresada. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*, y la planta ha sido transformada establemente con el vector mostrado en la Figura 8.

En otras realizaciones de la invención, la casete de expresión de ARNi se diseña para suprimir la expresión del polipéptido de XylT de la SEC ID N°: 6 o la SEC ID N°: 21, una variante biológicamente activa del polipéptido de XylT de SEC ID N°: 6 o la SEC ID N°: 21, o un polipéptido de XylT codificado por una secuencia que tiene al menos un 75% de identidad de secuencia con respecto a la SEC ID N°: 4, la SEC ID N°: 5, la SEC ID N°: 19 o la SEC ID N°: 20, por ejemplo, al menos el 75%, al menos el 80%, al menos el 85%, al menos el 90% o al menos el 95% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 4, la SEC ID N°: 5, la SEC ID N°: 19 o la SEC ID N°: 20. De esta manera, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo de la secuencia génica de XylT, en donde el fragmento directo comprende 318-1052 nt de la SEC ID N°: 4; una secuencia espaciadora que comprende aproximadamente de 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso de la secuencia génica de XylT, en donde el fragmento inverso comprende el complemento (es decir, la versión antisentido) de los nt 318-1052 de la SEC ID N°: 4. En una realización de este tipo, la secuencia espaciadora se representa mediante los nt 1053-1599 de la SEC ID N°: 4, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi correspondiente a la secuencia codificadora para la estructura de ARNhp es 2015 nt. Transformar establemente una planta con una construcción de nucleótidos que comprende dicha casete de expresión de ARNi, por ejemplo, el vector mostrado en la Figura 9, inhibe de manera efectiva la expresión de la FucT dentro de las células vegetales de la planta en la que se expresa la estructura de ARNhp. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*, y la planta ha sido transformada establemente con el vector mostrado en la Figura 9.

En otras realizaciones, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo de la secuencia génica de XylT, en donde el fragmento directo comprende los nt 1-734 de la SEC ID N°: 19; una secuencia espaciadora que comprende de aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso de la secuencia génica de XylT, en donde el fragmento inverso comprende el complemento (es decir, la versión antisentido) de los nt 1-734 de la SEC ID N°: 19. En dicha realización, la secuencia espaciadora está representada por los nt 735-1282 de la SEC ID N°: 19, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi que corresponde a la secuencia codificadora para la estructura de ARNhp es de 2015 nt. Transformar establemente una planta con una construcción de nucleótidos que comprende dicha casete de expresión de ARNi, por ejemplo, el vector mostrado en la Figura 9, inhibe de manera efectiva la expresión de FucT dentro de las células vegetales de la planta en las que se expresa la estructura de ARNhp. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*, y la planta se ha transformado establemente con el vector mostrado en la Figura 9.

En otras realizaciones adicionales, se pueden diseñar estructuras de ARNhp de mayor tamaño de tal modo que las secuencias de antisentido y sentido están en orientaciones opuestas. De este modo, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés, la secuencia génica de longitud completa de FucT o de XylT en la orientación antisentido, y un fragmento directo de la secuencia génica de FucT o XylT que comprende la mitad 3' de la secuencia en la orientación sentido (véase la Figura 28, diseño 1). En este tipo de construcción, la mitad 3' de la secuencia forma el tallo de bases emparejadas (es decir, de cadena doble) del ARNhp, y la mitad 5' de la secuencia actúa como secuencia espaciadora. Sin pretender establecer ninguna teoría o mecanismo, la región 3' de la secuencia de FucT o XylT se elige para formar la región de cadena doble del ARNhp para esta construcción debido

a que la región está relativamente conservada entre las especies vegetales en comparación con la región 5'.

En dicha realización, la casete de expresión de ARNi se diseña para suprimir la expresión del polipéptido de FucT de la SEC ID N°: 3, una variante biológicamente activa del polipéptido de FucT de SEC ID N°: 3, o un polipéptido de FucT codificado por una secuencia que tiene al menos un 75% de identidad de secuencia con respecto a la SEC ID N°: 1 o la SEC ID N°: 2, por ejemplo, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90% o al menos un 95% de identidad de secuencia con la secuencia de la SEC ID N°: 1 o la SEC ID N°: 2. De esta manera, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; los nucleótidos de 1-1865 de la SEC ID N°: 1 en orientación antisentido, y los nucleótidos de 950-1865 de la SEC ID N°: 1 en orientación sentido. Transformar establemente una planta con una construcción de nucleótidos que comprende esta casete de expresión de ARNi inhibe de manera efectiva la expresión de FucT dentro las células vegetales de la planta en las que se expresa la estructura de ARNhp. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*.

En otra realización de este tipo, la casete de expresión de ARNi se diseña para suprimir la expresión del polipéptido de XylT de SEC ID N°: 6 o la SEC ID N°: 21, una variante biológicamente activa del polipéptido de XylT de SEC ID N°: 6 o SEC ID N°: 21, o un polipéptido codificado por una secuencia que tiene al menos un 75% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 4, SEC ID N°: 5, SEC ID N°: 19 o SEC ID N°: 20, por ejemplo, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90% o al menos un 95% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 5, SEC ID N°: 19 o SEC ID N°: 20. De esta manera, la cadena sentido de la casete de expresión de ARNi se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés, los nucleótidos 1-1860 de la SEC ID N°: 4 en orientación antisentido, y los nucleótidos 950-1860 de la SEC ID N°: 4 en la orientación sentido. Transformar establemente una planta con una construcción de nucleótidos que comprende esta casete de expresión de ARNi inhibe de manera efectiva la expresión de XylT dentro las células vegetales de la planta en las que se expresa la estructura de ARNhp. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*.

Cuando se desea la supresión de ambas proteínas, FucT y XylT, se puede lograr introduciendo dichas casetes de expresión de ARNi de gen individual en la planta en un evento de transformación individual, por ejemplo, ensamblando dichas casetes de expresión de ARNi de gen individual dentro de un único vector de transformación, por ejemplo, un vector similar al mostrado en la Figura 11, o como eventos de co-transformación separados, por ejemplo, ensamblando dichas casetes de expresión de ARNi de gen individual dentro de dos vectores de transformación, por ejemplo, vectores similares a los mostrados en las Figuras 8 y 9, usando cualquier método de transformación adecuado conocido en la técnica, que incluyen, aunque sin limitación, los métodos de transformación descritos en otra parte de la presente memoria.

Alternativamente, la supresión de ambas proteínas, FucT y XylT se puede lograr introduciendo en la planta superior de interés una casete de expresión de ARNi quimérico como se ha indicado anteriormente en la presente invención. Por tanto, en algunas realizaciones de la invención, la cadena sentido de una casete de expresión de ARNi quimérico se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo quimérico, que comprende de aproximadamente 500 a aproximadamente 650 nucleótidos (nt) de una cadena sentido para la FucT y de aproximadamente 500 a aproximadamente 650 nt de una cadena sentido para la XylT, en donde la secuencia de FucT y la secuencia de XylT pueden estar en cualquier orden; una secuencia espaciadora que comprende de aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia; y un fragmento inverso del fragmento directo quimérico, en donde el fragmento inverso comprende la secuencia antisentido complementaria al respectivo fragmento directo quimérico.

Como se ha indicado previamente para las casetes de expresión de ARNi individuales, se reconoce que la secuencia individual de FucT o de XylT dentro del fragmento directo quimérico puede comprender una secuencia de nucleótido que es un 100% idéntica a la correspondiente porción de la cadena sentido de la secuencia génica diana de FucT y XylT, respectivamente, o como alternativa, puede comprender una secuencia que comparte al menos un 90%, al menos un 95% o al menos un 98% de identidad de secuencia con respecto a la correspondiente porción de la cadena sentido correspondiente al gen diana de FucT o XylT que va a ser silenciado. De un modo similar, se reconoce que el fragmento inverso no tiene que compartir una identidad de secuencia del 100% con el complemento del fragmento directo quimérico; sino que más bien debe tener una longitud suficiente y una complementariedad suficiente con la secuencia de fragmento directo quimérico, tal como se ha definido anteriormente en la presente memoria, de tal modo que cuando la molécula de ARN inhibidor se exprese, las regiones transcritas correspondientes al fragmento directo quimérico y al fragmento inverso se hibridarán para formar el tallo de bases emparejadas (es decir, la porción de cadena doble) de la estructura de ARNhp. Al diseñar dicha casete de expresión de ARNi, las longitudes del fragmento directo, de la secuencia espaciadora y de los fragmentos inversos se eligen de tal modo que la longitud combinada del polinucleótido que codifica la estructura de ARNhp es de aproximadamente 1200 a aproximadamente 3300 nt, de aproximadamente 1250 a aproximadamente 3300 nt, de aproximadamente 1300 a aproximadamente 3300 nt, de aproximadamente 1350 a aproximadamente 3300 nt, de aproximadamente 1400 a aproximadamente 3300 nt, de aproximadamente 1450 a aproximadamente 3300 nt, de aproximadamente 1500 a aproximadamente 3300 nt, de aproximadamente 2200 a aproximadamente 3100 nt, de aproximadamente

2250 a aproximadamente 2800 nt ó de aproximadamente 2500 a aproximadamente 2700 nt. En algunas realizaciones, la longitud combinada de la construcción de horquilla expresada es de aproximadamente 1200 nt, aproximadamente 1250 nt, aproximadamente 1300 nt, aproximadamente 1350 nt, aproximadamente 1400 nt, aproximadamente 1450 nt, aproximadamente 1500 nt, aproximadamente 1800 nt, aproximadamente 2200 nt, aproximadamente 2250 nt, aproximadamente 2300 nt, aproximadamente 2350 nt, aproximadamente 2400 nt, aproximadamente 2450 nt, aproximadamente 2500 nt, aproximadamente 2550 nt, aproximadamente 2600 nt, aproximadamente 2650 nt, aproximadamente 2700 nt, aproximadamente 2750 nt, aproximadamente 2800 nt, aproximadamente 2850 nt, aproximadamente 2900 nt, aproximadamente 2950 nt, aproximadamente 3000 nt, aproximadamente 3050 nt, aproximadamente 3100 nt, aproximadamente 3150 nt, aproximadamente 3200 nt, aproximadamente 3250 nt, aproximadamente 3300 nt, o cualquier longitud de este tipo entre aproximadamente 1200 nt y aproximadamente 3300 nt.

En algunas realizaciones, el fragmento quimérico directo comprende de aproximadamente 500 a aproximadamente 650 nt, por ejemplo, 500, 525, 550, 575, 600, 625 o 650 nt, de una cadena sentido para FucT, por ejemplo, de la cadena sentido establecida en la SEC ID N°: 1 ó 2, y de aproximadamente 500 a aproximadamente 650 nt, por ejemplo, 500, 525, 550, 575, 600, 625 ó 650 nt, de una cadena sentido para XylT, por ejemplo, de la cadena sentido establecida en la SEC ID N°: 4, 5, 19 ó 20, en donde la secuencia de FucT y XylT pueden estar condensada una con otra; la secuencia espaciadora comprende de aproximadamente 100 a aproximadamente 700 nt, por ejemplo, 100, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 500, 525, 550, 575, 600, 625, 650, 675 ó 700 nt de cualquier secuencia de interés; y el fragmento inverso comprende la cadena antisentido correspondiente a la secuencia de fragmento directo quimérico, o una secuencia que tenga una longitud suficiente y una complementariedad suficiente con la secuencia del fragmento directo quimérico.

Tal como se ha indicado anteriormente para las casetes de expresión de ARNi, la secuencia espaciadora puede ser cualquier secuencia que tenga una homología insuficiente con el gen diana, es decir, FucT o XylT, y una homología insuficiente consigo misma de tal modo que la porción de la molécula de ARN inhibidora expresada correspondiente a la región espaciadora no puede auto-hibridarse, y por tanto forma el lazo de la estructura de ARNhp. En algunas realizaciones, la secuencia espaciadora comprende un intrón, y por tanto la molécula de ARN inhibidora expresada forma un ARNihp tal como se ha indicado en la presente memoria anteriormente. En otras realizaciones, la secuencia espaciadora comprende una porción a la cadena sentido correspondiente al gen de FucT o XylT que va a silenciarse, por ejemplo, una porción de la cadena sentido establecida en la SEC ID N°: 1 ó 2 (FucT) o la SEC ID N°: 4, 5, 19 ó 20 (XylT). En una realización, el fragmento directo quimérico comprende la secuencia de FucT y XylT fusionadas en ese orden, y la secuencia espaciadora comprende una porción de la cadena sentido de XylT inmediatamente por debajo de la secuencia de XylT contenida dentro del fragmento directo quimérico. En otra realización, el fragmento directo quimérico comprende la secuencia de XylT y FucT fusionada en ese orden, y la secuencia espaciadora comprende una porción de la cadena sentido de FucT inmediatamente por debajo de la secuencia de FucT contenida dentro del fragmento directo quimérico.

En algunas realizaciones, la casete de expresión de ARNi quimérico se diseña para suprimir la expresión del polipéptido de FucT de SEC ID N°: 3, una variante biológicamente activa de la polipéptido de FucT de SEC ID N°: 3, o un polipéptido de FucT codificado por una secuencia que tiene al menos un 75% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 1 o la SEC ID N°: 2, por ejemplo, al menos un 75%, al menos 80%, al menos un 85%, al menos un 90% o al menos un 95% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 1 o la SEC ID N°: 2, y para suprimir la expresión del polipéptido de XylT de la SEC ID N°: 6 o SEC ID N°: 21, una variante biológicamente activa del polipéptido de XylT de SEC ID N°: 6 o SEC ID N°: 21, o un polipéptido de XylT codificado por una secuencia que tiene al menos un 75% de identidad con respecto a la secuencia de la SEC ID N°: 4, SEC ID N°: 5, SEC ID N°: 19 o SEC ID N°: 20, por ejemplo, al menos un 75%, al menos un 80%, al menos un 85%, al menos un 90% o al menos un 95% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 4, SEC ID N°: 5, SEC ID N°: 19 o SEC ID N°: 20. Para algunas de estas realizaciones, la secuencia de FucT dentro del fragmento directo quimérico se elige de tal modo que corresponda a los nt 700 a 1400 de la SEC ID N°: 1 o la SEC ID N°: 2, y/o la secuencia de XylT dentro del fragmento directo quimérico se elige de tal modo que corresponda a los nt 700 a 1400 de la SEC ID N°: 4 o la SEC ID N°: 5, o los nt 383 a 1083 de la SEC ID N°: 19 ó 20. Sin pretender establecer una teoría, se cree que esta región (particularmente para FucT) es relativamente conservada entre las diferentes especies de plantas, y por lo tanto es una diana potencialmente buena.

En otras realizaciones, la cadena sentido de la casete de expresión de ARNi quimérico se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo quimérico que comprende los nt 254-855 de la SEC ID N°: 1 (secuencia de FucT) y los nt 318-943 de la SEC ID N°: 4 (secuencia de XylT); una secuencia espaciadora que comprende aproximadamente de 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso que comprende el complemento (es decir, la versión antisentido) del fragmento quimérico directo, es decir, que comprende el complemento de los nt 318-943 de la SEC ID N°: 4 y el complemento de los nt 254-855 de la SEC ID N°: 1. En una realización particular, la secuencia espaciadora dentro de esta casete de expresión de ARNi quimérico se representa mediante los nt 944-1443 de la SEC ID N°: 4, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi correspondiente a la secuencia codificadora correspondiente a la estructura de ARNhp es de 2956 nt.

En otra realización de este tipo, la cadena sentido de la casete de expresión de ARNi quimérico se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo quimérico que comprende los nt 318-943 de la SEC ID N°: 4 (secuencia de XylIT) y los nt 254-855 de la SEC ID N°: 1 (secuencia de FucT); una secuencia espaciadora que comprende de aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso que comprende el complemento (es decir, la versión antisentido) del fragmento directo quimérico, es decir, que comprende el complemento de los nt 254-855 de la SEC ID N°: 1 y el complemento de los nt 318-943 de la SEC ID N°: 4. En una realización particular, la secuencia espaciadora dentro de esta casete de expresión de ARNi quimérico se representa por los nt 856-1355 de la SEC ID N°: 1, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi correspondiente a la secuencia codificadora para la estructura de ARNhp es de 2956 nt.

En otras realizaciones adicionales, la cadena sentido de la casete de expresión de ARNi quimérico se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo quimérico que comprende los nt 254-855 de la SEC ID N°: 1 (secuencia de FucT) y los nt 1-626 de la SEC ID N°: 19 (secuencia de XylIT); una secuencia espaciadora que comprende de aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia como la indicada anteriormente; y un fragmento inverso que comprende el complemento (es decir, la versión antisentido) del fragmento directo quimérico, es decir, que comprende el complemento de los nt 1-626 de la SEC ID N°: 19 y el complemento de los nt 254-855 de la SEC ID N°: 1. En una realización particular, la secuencia espaciadora dentro de esta casete de expresión de ARNi quimérico se representa mediante los nt 627-1126 de la SEC ID N°: 19, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi que corresponde a la secuencia codificadora para la estructura de ARNhp es de 2956 nt.

En otra realización de este tipo, la cadena sentido de la casete de expresión de ARNi quimérico se diseña para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor de interés; un fragmento directo quimérico que comprende los nt 1-626 de la SEC ID N°: 19 (secuencia de XylIT) y los nt 254-855 de la SEC ID N°: 1 (secuencia de FucT); una secuencia espaciadora que comprende de aproximadamente 100 a aproximadamente 700 nt de cualquier secuencia como se ha indicado anteriormente; y un fragmento inverso que comprende el complemento (es decir, la versión antisentido) del fragmento directo quimérico, es decir, que comprende el complemento de los nt 254-855 de la SEC ID N°: 1 y el complemento de los nt 1-626 de la SEC ID N°: 19. En una realización particular, la secuencia espaciadora dentro de esta casete de expresión de ARNi quimérico está representada por los nt 856-1355 de la SEC ID N°: 1, y la longitud total de dicha porción de la cadena sentido de la casete de expresión de ARNi correspondiente a la secuencia codificadora para la estructura de ARNhp es de 2956 nt.

Transformar establemente una planta con una construcción de nucleótidos que comprende una casete de expresión de ARNi descrita en la presente memoria, por ejemplo, la transformación estable con el vector mostrado en la Figura 10, inhibe de manera efectiva la expresión de ambas, FucT y XylIT, en las células vegetales de la planta en la que se expresa la estructura de ARNhp. En una realización, la planta de interés es un miembro de la familia de las lentejas de agua, por ejemplo, un miembro de las *Lemnaceae*, y la planta ha sido transformada establemente con el vector mostrado en la Figura 10.

Se reconoce que la planta puede transformarse establemente con al menos dos de estas casetes de expresión de ARNi quimérico para proporcionar un silenciamiento génico muy eficiente de las proteínas de FucT y XylIT, que incluye el silenciamiento de cualquier isoforma de estas dos proteínas. Véase, por ejemplo, las dos orientaciones proporcionadas en el "diseño posible 2" de la Figura 28. De esta manera, la planta puede ser transformada establemente con una primera casete de expresión de ARNi quimérico en donde el fragmento directo quimérico comprende la secuencia de FucT y XylIT fusionada en ese orden, y la secuencia espaciadora comprende una porción de la cadena sentido de XylIT inmediatamente por debajo de la secuencia de XylIT contenida dentro del fragmento directo quimérico; y con una segunda casete de expresión de ARNi quimérico en donde el fragmento director quimérico comprende la secuencia de XylIT y FucT fusionada en ese orden, y la secuencia espaciadora comprende una porción de la cadena sentido de la FucT inmediatamente por debajo de la secuencia de FucT contenida dentro del fragmento directo quimérico.

El promotor ligado operativamente dentro de cualquiera de las casetes de expresión de ARNi que codifican grandes estructuras de ARNhp, o grandes estructuras de ARNhp puede ser cualquier promotor de interés que proporcione la expresión del polinucleótido inhibidor ligado operativamente dentro de la planta de interés, que incluye uno de los promotores descritos en la presente memoria más adelante. La región reguladora puede comprender elementos reguladores adicionales que potencian la expresión del polinucleótido inhibidor, que incluyen, aunque sin limitación, las secuencias líder 5' y las secuencias líder 5' más intrones vegetales discutidas en la presente memoria más adelante.

En otras realizaciones adicionales, la casete de expresión de ARNi puede diseñarse para proporcionar la expresión de estructuras de ARNhp pequeñas que tienen una región de tallo de bases emparejadas que comprende aproximadamente 200 pares base o menos. La expresión de la estructura pequeña de ARNhp es dirigida preferiblemente por un promotor reconocido por ARN polimerasa III dependiente de ADN. Véase, por ejemplo, la

Solicitud de Patente de EE.UU. N°: 2004023101.

De esta manera, la casete de expresión de ARNi se diseña de tal modo que la región de ADN transcrita codifica una molécula de ARN que comprende una región de nucleótidos sentido y antisentido, en donde la secuencia de nucleótidos sentido comprende aproximadamente 19 nucleótidos contiguos que tienen de aproximadamente 90% a aproximadamente 100% de identidad de secuencia con respecto a una secuencia de nucleótidos de aproximadamente 19 nucleótidos contiguos procedentes del ARN transcrito del gen de interés y en donde la secuencia de nucleótidos antisentido comprende aproximadamente 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con el complemento de una secuencia de nucleótidos de aproximadamente 19 nucleótidos contiguos de la secuencia sentido. Las secuencias de nucleótidos sentido y antisentido de la molécula de ARN deberían ser capaces de formar un región de tallo de bases emparejadas (es decir, de cadena doble) de ARN de aproximadamente 19 a aproximadamente 200 nucleótidos, alternativamente de aproximadamente 21 a aproximadamente 90 ó 100 nucleótidos, o alternativamente de aproximadamente 40 a aproximadamente 50 nucleótidos de longitud. Sin embargo, la longitud de la región de tallo de bases emparejadas de la molécula de ARN también puede tener una longitud de aproximadamente 30, aproximadamente 60, aproximadamente 70 o aproximadamente 80 nucleótidos. Cuando la región de tallo de bases emparejadas de la molécula de ARN es más grande de 19 nucleótidos, sólo hay un requisito de que haya al menos una región de cadena doble de aproximadamente 19 nucleótidos (en donde puede haber aproximadamente una discordancia entre la región sentido y la antisentido) cuya cadena sentido es "idéntica" (permitiendo una discordancia) con 19 nucleótidos consecutivos del polinucleótido diana de FucT o XylT de interés. La región de ADN transcrita de este tipo de casete de expresión de ARNi puede comprender una secuencia espaciadora posicionada entre la región de nucleótidos codificador sentido y antisentido. La secuencia espaciadora no está relacionada con el polinucleótido de FucT o XylT diana, y puede oscilar en longitud entre 3 y aproximadamente 100 nucleótidos, o alternativamente entre aproximadamente 6 y aproximadamente 40 nucleótidos. Este tipo de casete de expresión de ARNi también puede comprender una secuencia terminadora reconocida por la ARN polimerasa III, siendo la secuencia una tira oligo dT, posicionada por debajo de la región de nucleótidos codificadora antisentido de la casete. Por "tira oligo dT" se entiende una tira de residuos T consecutivos. Debería comprender al menos 4 residuos T, pero obviamente puede contener más residuos T.

Se reconoce que al diseñar un ARNhp corto, los fragmentos de la secuencia génica diana (es decir, los fragmentos de secuencia génica de FucT o XylT9 y cualquier secuencia espaciadora que vayan a ser incluidas dentro de la porción codificadora de ARNhp de la casete de expresión de ARNi se eligen para evitar secuencias ricas en GC, particularmente aquellas con tres G/C's consecutivas, y para evitar la presencia de cuatro o más T's o A's consecutivos, ya que la cuerda "TTTT..." actúa como secuencia terminadora reconocida por la ARN polimerasa III.

Por tanto, cuando se desea un silenciamiento de gen con un ARNhp corto, la casete de expresión de ARNi puede diseñarse para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor reconocido por una ARN polimerasa III dependiente de ADN de la célula vegetal, tal como se define en la presente memoria más adelante; un fragmento de ADN que comprende una secuencia de nucleótidos sentido y antisentido, en donde la secuencia de nucleótidos sentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto a una secuencia de nucleótidos de al menos 19 nucleótidos contiguos procedentes de la cadena sentido del gen de FucT o XylT de interés, y en donde la secuencia de nucleótidos antisentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto al complemento de una secuencia de nucleótidos de al menos 19 nucleótidos contiguos de la secuencia sentido, en donde las secuencias de nucleótidos sentido y antisentido son capaces de formar un ARN de cadena doble de aproximadamente 19 a aproximadamente 200 nucleótidos de longitud; y una tira oligo dT que comprende al menos 4 residuos T consecutivos.

En algunas realizaciones de la invención, la casete de expresión de ARNi se diseña para expresar un ARNhp pequeño que suprima la expresión del polipéptido de FucT de la SEC ID N°: 3, una variante biológicamente activa del polipéptido de FucT de la SEC ID N°: 3, o un polipéptido de FucT codificado por una secuencia que tiene al menos un 90% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 1 ó la SEC ID N°: 2. De esta manera, la casete de expresión de ARNi se puede diseñar para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor reconocido por una ARN polimerasa III dependiente de ADN de la célula vegetal, tal como se define en la presente memoria más adelante; un fragmento de ADN que comprende una secuencia de nucleótidos sentido y antisentido, en donde la secuencia de nucleótidos sentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto a una secuencia de nucleótidos de al menos 19 nucleótidos contiguos de la SEC ID N°: 1, y en donde la secuencia de nucleótidos antisentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto al complemento de una secuencia de nucleótidos de al menos 19 nucleótidos contiguos de la secuencia sentido, en donde las secuencias de nucleótidos sentido y antisentido son capaces de formar un ARN de cadena doble de aproximadamente 19 a aproximadamente 200 nucleótidos de longitud; y una tira oligo dT que comprende al menos 4 residuos T consecutivos.

En otras realizaciones de la invención, la casete de expresión de ARNi se diseña para expresar un ARNhp pequeño

que suprime la expresión del polipéptido de XylT de la SEC ID N°: 6 o la SEC ID N°: 21, una variante biológicamente activa del polipéptido de XylT de la SEC ID N°: 6 o la SEC ID N°: 21, o un polipéptido de XylT codificado por una secuencia que tiene al menos un 90% de identidad de secuencia con respecto a la secuencia de la SEC ID N°: 4, la SEC ID N°: 5, la SEC ID N°: 19 o la SEC ID N°: 20. De esta manera, la casete de expresión de ARNi se puede diseñar para comprender en la dirección 5'-a-3' los siguientes elementos ligados operativamente: un promotor reconocido por una ARN polimerasa III dependiente de ADN de la célula vegetal, tal como se define en la presente memoria más adelante; un fragmento de ADN que comprende la secuencia de nucleótidos sentido y antisentido, en donde la secuencia de nucleótidos sentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto a la secuencia de nucleótidos de al menos 19 nucleótidos contiguos de la SEC ID N°: 4, y en donde la secuencia de nucleótidos antisentido comprende al menos 19 nucleótidos contiguos que tienen de aproximadamente un 90% a aproximadamente un 100% de identidad de secuencia con respecto al complemento de una secuencia de nucleótidos de al menos 19 nucleótidos contiguos de la secuencia sentido, en donde las secuencias de nucleótidos sentido y antisentido son capaces de formar un ARN de cadena doble de aproximadamente 19 a aproximadamente 200 nucleótidos de longitud; y una tira oligo dT que comprende al menos 4 residuos T consecutivos.

ARN interferente pequeño o Micro ARN:

En algunas realizaciones de la invención, la inhibición de la expresión de FucT o XylT, o de ambas, se puede obtener mediante interferencia de ARN por expresión de un gen que codifica un micro ARN (miARN). Los miARNs son agentes reguladores que consiste en aproximadamente 22 ribonucleótidos. Los miARN son altamente efectivos para inhibir la expresión.

Para la interferencia de miARN, la casete de expresión se diseña para expresar una molécula de ARN que se modela sobre un gen de miARN endógeno. El gen de miARN codifica un ARN que forma una estructura de horquilla que contiene una secuencia de 22 nucleótidos que es complementaria con otro gen endógeno (secuencia diana). Para la supresión de la expresión de FucT o XylT, se selecciona la secuencia de 22 nucleótidos a partir de una secuencia de transcrito de FucT o XylT y contiene 22 nucleótidos de dicha secuencia de FucT o XylT en orientación sentido y 21 nucleótidos de una secuencia antisentido correspondiente que es complementaria a la secuencia sentido. Las moléculas de miARN son altamente eficientes para inhibir la expresión de genes endógenos, y la interferencia de ARN que inducen es heredada por las generaciones posteriores de plantas.

Por tanto, la inhibición de la expresión de FucT y/o XylT en una planta superior de interés se puede lograr mediante cualquiera de los anteriores métodos con el fin de alterar el modelo de *N*-glicosilación de glicoproteínas endógenas y heterólogas producidas dentro de dicha planta de tal modo que dichas glicoproteínas comprenden *N*-glicanos complejos que tienen una reducción en la cantidad de residuos de xilosa β 1,2-ligados y/o residuos de fucosa α 1,3-ligados. La extensión de la reducción de la unión del residuo de xilosa β 1,2-ligado y/o el residuo de fucosa α 1,3-ligado con los *N*-glicanos de glicoproteínas está gobernada por el grado de inhibición de la expresión de las respectivas enzimas de XylT y FucT.

En algunas realizaciones de la invención, las glicoproteínas recombinantes producidas en un hospedante vegetal que es transformado establemente usando los métodos descritos en la presente memoria para afectar a la expresión de XylT en glicanos *N*-ligados que comprende menos del 50%, menos del 40%, menos del 30% de los residuos de xilosa β 1,2-ligados presentes en los respectivos glicanos *N*-ligados de las glicoproteínas producidas en un hospedante vegetal que no ha sido modificado genéticamente para inhibir la expresión de la enzima XylT y las isoformas de la misma. En otras realizaciones, dichas glicoproteínas recombinantes tienen glicanos *N*-ligados que comprenden menos de un 25%, menos de un 20%, menos de un 15%, menos de un 10%, menos de un 5% o menos de un 1% de los residuos de xilosa β 1,2-ligados presentes en los respectivos glicanos *N*-ligados o glicoproteínas producidas en un hospedante vegetal que no ha sido modificado genéticamente para inhibir la expresión de la enzima de XylT y las isoformas de la misma. En otras realizaciones, los métodos de la invención proporcionan un silenciamiento completo del gen de XylT y de cualquier isoforma del mismo dentro de la planta transformada establemente, de tal modo que las glicoproteínas recombinantes producidas dentro de la planta tienen glicanos *N*-ligados que están desprovistos de residuos de xilosa β 1,2-ligados.

De un modo similar, cuando un hospedante vegetal ha sido transformado establemente usando los métodos descritos en la presente memoria para afectar a la expresión de FucT, las glicoproteínas recombinantes producidas dentro de la planta tienen glicanos *N*-ligados que comprenden menos del 50%, menos del 40%, menos del 30% de los residuos de fucosa α 1,3-ligados presentes en los respectivos glicanos *N*-ligados de las glicoproteínas producidas en un hospedante vegetal que no ha sido modificado genéticamente para inhibir la expresión de la enzima FucT y las isoformas de la misma. En otras realizaciones, estas glicoproteínas recombinantes tienen glicanos *N*-ligados que comprenden menos del 25%, menos del 20%, menos del 15%, menos del 10%, menos del 5% o menos del 1% de los residuos de fucosa α 1,3-ligados presentes en los respectivos glicanos *N*-ligados de las glicoproteínas producidas en una planta hospedante que no ha sido modificada genéticamente para inhibir la expresión de la enzima FucT y de las isoformas de la misma. En otras realizaciones adicionales, los métodos de la invención proporcionan un silenciamiento completo del gen de FucT y de cualquiera de sus isoformas dentro de la planta transformada establemente, de tal modo que las glicoproteínas recombinantes producidas dentro de la planta tienen glicanos *N*-ligados que están desprovistos de residuos de fucosa α 1,3-ligados.

Cuando un hospedante vegetal ha sido transformado establemente usando los métodos descritos en la presente memoria para afectar a la expresión de ambas enzimas, XylT y FucT, y cualquiera de sus isoformas, las glicoproteínas recombinantes producidas dentro de la planta tienen glicanos N-ligados que comprenden menos del 50%, menos del 40%, menos del 30% de los residuos de xilosa β 1,2-ligados y menos del 50%, menos del 40%, menos del 30% de los residuos de fucosa α 1,3-ligados presentes en los respectivos glicanos N-ligados de glicoproteínas producidas en un hospedante vegetal que no ha sido modificado genéticamente para inhibir la expresión de las enzimas XylT y FucT y las isoformas de los mismos. En otras realizaciones, estas glicoproteínas recombinantes tienen glicanos N-ligados que comprenden menos del 25%, menos del 20%, menos del 15%, menos del 10%, menos del 5% o menos del 1% de los residuos de xilosa β 1,2-ligados y el 25%, menos del 20%, menos del 15%, menos del 10%, menos del 5% o menos del 1% de los residuos de fucosa α 1,3-ligados presentes en los respectivos glicanos N-ligados de glicoproteínas producidas en un hospedante vegetal que no ha sido modificado genéticamente para inhibir la expresión de las enzimas XylT y FucT y las isoformas de las mismas. En otras realizaciones adicionales, los métodos de la invención proporcionan el silenciamiento completo del gen de XylT y FucT y de cualquiera de sus isoformas dentro de la planta transformada establemente, de tal modo que las glicoproteínas recombinantes producidas dentro de la planta tienen glicanos N-ligados que están desprovistos de residuos de xilosa β 1,2-ligados y de residuos de fucosa α 1,3-ligados.

En algunas realizaciones de la presente invención, un hospedante vegetal que ha sido transformado establemente usando los métodos descritos en la presente memoria para afectar a la expresión de ambas enzimas, XylT y FucT, y a cualquiera de sus isoformas, es capaz de producir glicoproteínas recombinantes en las que los glicanos N-ligados son sustancialmente homogéneos. Por "sustancialmente homogéneos" se entiende que el perfil de glicosilación refleja la presencia de un único pico principal correspondiente a la especie de N-glicano deseada, más particularmente, la especie de glicano G0, en donde al menos el 90% de las estructuras de N-glicano presentes en dichas glicoproteínas son de la especie de glicano G0.

Los métodos para monitorizar los cambios en el modelo de N-glicosilación de glicoproteínas, también denominados perfiles de glicosilación, son bien conocidos en la técnica e incluyen, aunque sin limitación, la espectrometría de masas de tiempo de vuelo de ionización por desorción de láser asistida por matriz (MALDI-TOF), por ejemplo, usando el ensayo MALDI-TOF descrito en el Ejemplo 3 de la presente memoria, la espectrometría de masas acoplada a cromatografía de líquidos (LC-MS), la cromatografía de gases, la cromatografía de intercambio aniónico, la cromatografía por exclusión de tamaños, la electroforesis capilar y la electroforesis de gel capilar, el marcado con fluorescencia y la detección mediante cromatografía de líquidos de alta resolución (HPLC) y QTOF, y otras similares. De esta manera, los cambios en el modelo de N-glicosilación debidos a la inhibición de la expresión o la función de XylT y/o FucT en una planta transformada establemente de la invención se pueden monitorizar sometiendo una muestra (por ejemplo una muestra de tejido de hoja) obtenida de una planta transformada establemente a un análisis total de N-glicanos mediante espectrometría de masas MALDI-TOF, y comparando los resultados con los obtenidos para una muestra comparable procedente de una planta de control, en donde la planta de control no ha sido modificada genéticamente para inhibir la expresión o la función de XylT y/o FucT. Una reducción de la cantidad de residuos de xilosa y/o fucosa en los N-glicanos se puede monitorizar mediante una reducción de la masa de los respectivos picos. Véase, por ejemplo, Strasser et al. (2004) FEBS Letters 561: 132-136.

De forma similar, el perfil de glicosilación de cualquier glicoproteína producida recombinantemente se determina fácilmente usando técnicas estándares bien conocidas por los especialistas en la técnica. Véase, por ejemplo, la revisión proporcionada en Morelle y Michalski (2005) Curr. Anal. Chem. 1: 29-57. Así, las glicoproteínas que han sido producidas recombinantemente en un organismo hospedante, que incluyen un hospedante vegetal, se pueden analizar para determinar la relación de las estructuras particulares de glicanos N-ligados unidos a las mismas. De esta manera, una muestra que comprende una glicoproteína aislada producida recombinantemente puede ser sometida a una reacción enzimática o química para liberar las estructuras de glicano individuales de la glicoproteína. Después de esta etapa de desglicosilación, se puede llevar a cabo el análisis del perfil de glicosilación usando cualquiera de los ensayos analíticos descritos en la presente memoria anteriormente.

Los productos de glicoproteínas producidos recombinantemente habitualmente existen como una población diversa de glicofomas que portan entre una y varias docenas de glicanos diferentes en cantidades molares variables en los sitios de glicosilación con grados variables de ocupación de sitio. Dependiendo de la glicoproteína, las diferentes glicofomas pueden producir diferentes perfiles funcionales. Por tanto, en algunas realizaciones de la invención, es deseable determinar el perfil de glicosilación de la glicoproteína que tienen los N-glicanos intactos. Se puede usar cualquier procedimiento conocido en la técnica para determinar el perfil de glicosilación de una glicoproteína intacta, incluyen los métodos de espectrometría de masas indicados anteriormente y en los ejemplos presentados más adelante en la presente memoria.

Reduciendo o eliminado la expresión o la función de la fucosiltransferasa y/o la xilosiltransferasa del modo establecido en la presente memoria, tanto temporal como establemente, es posible producir una planta transgénica superior que tenga la capacidad de producir glicoproteínas que tengan un perfil de N-glicosilación con una heterogeneidad reducida respecto a la observada normalmente para glicoproteínas producidas por esta planta cuando la expresión o la función de dichas enzimas no ha sido alterada (es decir, la planta tiene la maquinaria de glicosilación nativa o natural). Cuando la expresión o la función de una o ambas de estas enzimas es reducida establemente o eliminada usando uno o más de los métodos descritos anteriormente en la presente memoria, la

reducción de la heterogeneidad del perfil de *N*-glicosilación de las glicoproteínas producidas por la planta transgénica se puede mantener de generación en generación de plantas, incluyendo reproducción asexual o sexual, y se puede determinar a través de condiciones de cultivo y con escalado en producción.

De esta manera, la presente invención proporciona un método para reducir la heterogeneidad del perfil de *N*-glicosilación de una glicoproteína producida en una planta superior, por ejemplo, una planta dicotiledónea o monocotiledónea, por ejemplo, una planta de lentejas de agua. El método comprende introducir en la planta una construcción de nucleótidos descrita en la presente memoria de tal modo que la expresión o la función de la fucosiltransferasa y/o la xilosiltransferasa es reducida o eliminada dentro de la planta. En algunas realizaciones de la invención, el método para reducir la heterogeneidad del perfil de *N*-glicosilación de una glicoproteína producida en una planta superior comprende introducir en la planta superior de interés al menos una construcción de nucleótidos descrita en la presente memoria anteriormente, en donde la(s) construcción(es) de nucleótidos para la supresión de la expresión de fucosiltransferasa y/o xilosiltransferasa en la planta, por ejemplo, usando uno o más de los métodos descritos en la presente memoria anteriormente.

Por "reducir la heterogeneidad del perfil de *N*-glicosilación" se entiende que el perfil de *N*-glicosilación se caracteriza por una reducción del número total de especies de *N*-glicano distintas que aparecen en el perfil. Por tanto, por ejemplo, en donde una glicoproteína producida por una planta superior que tiene la maquinaria de glicosilación nativa o natural (y por tanto que no ha sido modificada genéticamente para reducir o eliminar la expresión de fucosiltransferasa y xilosiltransferasa) produce una glicoproteína con un perfil de *N*-glicosilación que se caracteriza por la presencia de una mezcla de 5 especies de *N*-glicano, se pueden usar los métodos de la invención para reducir el número de especies de *N*-glicano que aparecen en el perfil de *N*-glicosilación. De esta manera, cuando se modifica genéticamente una planta superior en la manera establecida en la presente memoria para reducir o eliminar la expresión o la función de fucosiltransferasa y/o xilosiltransferasa, el perfil de *N*-glicosilación de esta glicoproteína se caracterizaría por una reducción del número de especies de *N*-glicano que aparecen en el perfil, por ejemplo, una mezcla de menos de 5 especies de *N*-glicano, por ejemplo, 4, 3 ó 2 especies de *N*-glicano, o incluso una única especie de *N*-glicano. Cuando la heterogeneidad del perfil de *N*-glicosilación se reduce de tal modo que el perfil se caracteriza por la presencia de una única especie de *N*-glicano predominante, el perfil de *N*-glicosilación sería sustancialmente homogéneo para dicha especie de *N*-glicano.

Los métodos para reducir la heterogeneidad del perfil de *N*-glicosilación de una glicoproteína producida en una planta superior da como resultado la glicoproteína producida que tiene un perfil de *N*-glicosilación que es sustancialmente homogéneo para la especie de glicano G0. Los métodos para reducir la heterogeneidad del perfil de *N*-glicosilación de una glicoproteína producida en una planta superior da como resultado la glicoproteína producida que tiene un perfil de *N*-glicosilación sustancialmente homogéneo, en donde al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos un 99% de la cantidad total de especies de *N*-glicano que aparecen en el perfil de *N*-glicosilación correspondiente a la glicoproteína se representa mediante la especie de glicano G0. En estas realizaciones, puede aparecer una cantidad traza de especies de *N*-glicano precursoras en el perfil de *N*-glicosilación como se indica en otra parte de la presente memoria, en donde cualquier especie de *N*-glicano precursora dada que esté presente en el perfil de *N*-glicosilación está presente en menos del 5%, preferiblemente menos del 4%, menos del 3%, menos del 2%, menos del 1% e incluso menos del 0,5% o incluso menos del 0,1% de la cantidad total de especies de *N*-glicano que aparecen en el perfil.

La glicoproteína puede ser una glicoproteína endógena de interés, o puede ser una glicoproteína heteróloga que es producida por la planta superior de interés, por ejemplo, una glicoproteína de mamífero, que incluye, por ejemplo, las glicoproteínas descritas en la presente memoria en otra parte. En algunas realizaciones, la glicoproteína es un anticuerpo monoclonal. En otras realizaciones, la glicoproteína se selecciona del grupo que consiste en un interferón, eritropoietina (EPO), activador de plasminógeno de tejido (tPA), plasminógeno, factor estimulante de colonia de granulocito-macrófago (GM-CSF) e inmunoglobulinas terapéuticas.

Usando los métodos de la presente invención, es posible mantener la heterogeneidad reducida dentro del perfil de *N*-glicosilación para una glicoproteína producida en la planta transgénica con un escalado de la producción, y por tanto la planta continua produciendo la glicoproteína de tal modo que el perfil de *N*-glicosilación se caracteriza por una reducción del número de especies de *N*-glicano del perfil. Mediante "escalado de la producción" o "aumento de la escala de producción" se pretende indicar un aumento de la cantidad de biomasa vegetal que está presente dentro de un sistema de cultivo (es decir, un recipiente de cultivo o un contenedor de cultivo dentro del cual se cultiva la planta) que está siendo usado para producir una proteína de interés, en este caso, una glicoproteína de interés. Por tanto, el escalado de la producción se produce, por ejemplo, cuando se aumenta la escala de producción desde una escala que es adecuada para propósitos de investigación hasta una que es adecuada para producción a escala piloto, y se aumenta más hasta una escala que es adecuada para producción comercial de la glicoproteína de interés. La planta transgénica es una lenteja de agua, que actúan como un hospedante para la producción recombinante de una glicoproteína, y la heterogeneidad reducida del perfil de *N*-glicosilación de la glicoproteína producida recombinantemente se mantiene con el aumento de la escala de producción, en donde la escala de producción aumenta al menos 300 veces, al menos 500 veces, al menos 700 veces, al menos 1.000 veces, al menos 1.500 veces o más sobre la biomasa inicial de partida. En algunas de estas realizaciones, la planta es una planta de lenteja de agua que produce recombinantemente una glicoproteína de interés, y la heterogeneidad reducida del perfil de *N*-glicosilación se mantiene con el aumento de la escala de producción, en donde la escala de

producción aumenta en al menos 2.000 veces, al menos 3.000 veces, al menos 4.000 veces, al menos 5.000 veces, al menos 6.000 veces, al menos 6.500 veces o más sobre la biomasa inicial de partida. En una realización de este tipo, la planta superior es una planta de lenteja de agua que produce recombinantemente una glicoproteína de interés, y la heterogeneidad reducida del perfil de *N*-glicosilación para dicha glicoproteína se mantiene al aumentar la escala de producción, en donde la escala de producción aumenta en al menos 7.000 veces, 8.000 veces, 9.000 veces, 10.000 veces, 12.500 veces, 15.000 veces, 17.500 veces, 20.000 veces, 23.000 veces, 26.000 veces o más sobre la biomasa inicial de partida.

Además, cuando la planta transgénica de interés va a mantenerse mediante cultivo clonal continuo, la línea transgénica resultante continúa produciendo glicoproteína que exhibe la heterogeneidad reducida dentro de su perfil de *N*-glicosilación. El cultivo clonal continuo se puede lograr usando cualquier método conocido en la técnica. En algunas realizaciones, el cultivo clonal continuo se logra tomando periódicamente una o más submuestras del cultivo de plantas y transfiriendo la(s) submuestra(s) a un medio de cultivo fresco para continuar con el cultivo. Por tanto, por ejemplo, en algunas realizaciones, la línea vegetal transgénica que se mantiene mediante cultivo clonal continuo se una línea de planta transgénica de lenteja de agua que ha sido modificada genéticamente para reducir o eliminar la expresión o la función de la fucosiltransferasa y/o la xilosiltransferasa. De esta manera, la heterogeneidad reducida del perfil de *N*-glicosilación de las glicoproteínas producidas en la línea vegetal transgénica se mantiene con el cultivo clonal continuo de la línea vegetal transgénica durante al menos 8 meses, al menos 10 meses, al menos 1 año, al menos 1,5 años, al menos 2 años, al menos 2,5 años, al menos 3 años, al menos 3,5 años, al menos 4 años, al menos 4,5 años, al menos 5 años, o más, y se puede mantener tanto tiempo como se mantenga la línea vegetal transgénica.

Polipéptidos y Glicoproteínas Heterólogos:

Las lentejas de agua, que sirven como sistemas de expresión para proteínas recombinantes para uso farmacéutico, que han sido transformadas establemente para producir glicoproteínas con un modelo de *N*-glicosilación alterado usando los métodos descritos en la presente memoria pueden modificarse genéticamente para producir cualquier proteína recombinante de interés. Cuando la proteína recombinante es una en la que se puede aplicar glicosilación post-traduccional, los métodos de la invención proporcionan de manera ventajosa un medio para producir dichas glicoproteínas con un modelo de *N*-glicosilación que refleje más estrechamente la de hospedantes mamíferos, particularmente un modelo de glicosilación que esté "humanizado". Los ejemplos de proteínas recombinantes de interés incluyen, aunque sin limitación, insulina, hormona del crecimiento, plasminógeno, α -interferón, β -interferón, β -glucocerebrosidasa, β -glucoronidasa, proteína de retinoblastoma, proteína p53, angiostatina, leptina, anticuerpos monoclonales y fragmentos de los mismos, citoquinas, por ejemplo, eritropoietina (EPO), factor estimulador de macrófago granulocito, activador de plasminógeno de tejido, factores de coagulación sanguínea, por ejemplo, Factor VII, Factor VIII, Factor IX y proteína C activada, 1-antitripsina alfa, receptores, hormonas, vacunas humanas, vacunas animales, péptidos y albúmina de suero. Además, las plantas superiores transgénicas de la invención son capaces de producir un producto de glicoproteína que tiene un perfil de glicosilación sustancialmente homogéneo para la especie de glicano G0, y que se caracteriza por su homogeneidad sustancial para la glicoforma G0. Esto da como resultado de forma ventajosa sistemas de expresión de hospedante vegetal que tienen un mejor consistencia de producción, así como un menor riesgo químico, de fabricación y de control (CMC, del inglés "Chemical, Manufacturing and Control") asociado a la producción de dichas composiciones de glicoproteínas.

Las composiciones de glicoproteínas producidas de acuerdo con los métodos de la presente invención pueden ser contenidas en una composición que comprende un vehículo farmacéuticamente aceptable. Dichas composiciones son útiles en un método para tratar en un sujeto una enfermedad o trastorno para el cual el tratamiento con la glicoproteína proporcionará un beneficio terapéutico. De esta manera, las glicoproteínas que han sido producidas en una planta, por ejemplo, una lenteja de agua, transformada establemente de acuerdo a los métodos de la presente invención, pueden administrarse a un sujeto que lo necesite.

Las composiciones de glicoproteínas comprenden glicanos *N*-ligados que predominantemente tienen la estructura de glicano G0. De esta manera, la presente invención describe composiciones de glicoproteínas que tienen perfiles de glicosilación que son "sustancialmente homogéneos" o "sustancialmente uniformes" o que tienen una "homogeneidad sustancial", tal como se ha definido anteriormente en la presente memoria. Por tanto, las composiciones de glicoproteína son sustancialmente homogéneas respecto a la especie de glicano G0, y por tanto tienen un perfil de glicosilación sustancialmente homogéneo, en donde al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos el 99% de la cantidad total de especies de *N*-glicano que aparecen en el perfil de glicosilación correspondiente a la composición están representadas por la especie de glicano G0, apareciendo en el perfil de glicosilación cantidades traza de especies de *N*-glicano precursoras, es decir, cualquier especie de *N*-glicano precursora que esté presente en el perfil de glicosilación lo está en menos del 5%, preferiblemente menos del 4%, menos del 3%, menos del 2%, menos del 1% e incluso menos del 0,5%, o incluso menos del 0,1% de la cantidad total de especies de *N*-glicano que aparecen en el perfil. Para dicha composición, una especie de *N*-glicano precursora representativa que aparece en su perfil de glicosilación serían las especies de *N*-glicano precursoras Man3GlcNAc2, MGn (GlcNAc1Man3GlcNAc2 en donde GlcNAc1 está unida al brazo de 1,3 manosa) y GnM (GlcNAc1Man3GlcNAc2, en donde GlcNAc1 está unida al brazo de 1,6-manosa) descritas anteriormente, en donde puede estar presente cualquiera de estas especies de *N*-glicano precursoras individualmente o cualquiera de sus combinaciones.

De esta manera, la invención se refiere a composiciones de glicoproteína “sustancialmente homogéneas” o “sustancialmente uniformes” o a composiciones de glicoproteínas que tienen una “homogeneidad sustancial” tal como se ha definido anteriormente en la presente memoria, en donde al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos el 99% de la glicoproteínas presentes en la composición están representadas por la glicofoma G0, en donde todos los sitios de glicosilación anticipados están ocupados por las especies de glicano G0, con una cantidad traza de precursores o glicofomas no deseadas presentes en la composición, es decir, las glicofomas precursoras representan menos del 5%, menos del 4%, menos del 3%, menos del 2%, menos del 1% o incluso menos del 0,5% o menos del 0,1% de las glicofomas totales presentes en la composición. En dicha composición, una glicofoma precursora representativa sería aquella en la que los sitios de glicosilación no están ocupados, y un ejemplo de glicofoma no deseada sería una glicofoma que tiene una mezcla de glicano G0 y especies de glicano G0X o G0XF3 unidas a sus sitios de glicosilación.

En algunas realizaciones de la invención, el hospedante vegetal comprende uno o más polinucleótidos que proporcionan la expresión de un anticuerpo que se une específicamente a una proteína de mamífero de interés, particularmente a una proteína humana de interés. Por tanto, en un aspecto, la invención proporciona métodos para producir anticuerpos monoclonales en plantas superiores, en donde los anticuerpos monoclonales tienen un modelo de *N*-glicosilación que refleja una reducción de la cantidad de residuos de xilosa β 1,2-ligados y residuos de fucosa α 1,3-ligados dentro de los glicanos *N*-ligados, y las composiciones que comprenden anticuerpos monoclonales recombinantes producidos usando hospedantes vegetales modificados genéticamente de la manera establecida en la presente memoria. El hospedante vegetal es un miembro de la familia de las lentejas de agua.

Los anticuerpos monoclonales están siendo usados de manera creciente como agentes terapéuticos para tratar enfermedades humanas, que incluyen, aunque sin limitación, cáncer y enfermedades que tienen un componente autoinmune o inflamatorio. Véase, por ejemplo, King (1999) *Curr. Opin. Drug Discovery Dev.* 2: 110-17; Vaswani y Hamilton (1998) *Ann. Allergy Asthma Immunol.* 81: 105-19; y Holliger y Hoogenboom, *Nat. Biotechnology* 16: 1015-16. Aunque algunos de estos anticuerpos tienen efectos terapéuticos que producen solo la unión de antígeno, por ejemplo anticuerpos que se unen a un receptor o ligando para prevenir interacciones ligando-receptor, otros anticuerpos necesitan funciones efectoras tales como el reclutamiento del sistema inmune para matar células diana con el objetivo de ser terapéuticamente activos. Véase, por ejemplo, Clynes et al. (2000) *Nat. Med.* 6: 443-46; Clynes et al. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95: 652-56, y Anderson et al. (1997) *Biochem. Soc. Trans.* 25: 705-8.

Las actividades de reconocimiento de antígeno y las funciones efectoras de los anticuerpos residen en porciones diferentes de la molécula del anticuerpo. La porción Fab' del anticuerpo proporciona actividad de reconocimiento de antígeno, mientras que la porción Fc proporciona funciones efectoras tales como la activación de células efectoras accesorias que incluyen células fagocíticas (macrófagos y neutrófilos), células asesinas naturales y mastocitos. Los anticuerpos se unen a las células a través de la región Fc, con un sitio de receptor Fc sobre la región Fc del anticuerpo que se une a un receptor Fc (FcR) sobre la célula. Existe un número de receptores Fc que son específicos para diferentes clases de anticuerpos, que incluyen IgG (receptores gamma), IgE (receptores eta), IgA (receptores alfa) e IgM (receptores mu). La unión del anticuerpo a los receptores Fc de la superficie celular activa una serie de respuestas biológicas importantes y diversas que incluyen el atrapamiento y la destrucción de partículas recubiertas de anticuerpo, la eliminación de complejos inmunes, la lisis de células diana recubiertas de anticuerpo por células asesinas (denominada citotoxicidad mediada por célula dependiente de anticuerpo, o ADCC), el inicio de citotoxicidad dependiente de complemento (CDC), la liberación de mediadores inflamatorios y el control de la producción de inmunoglobulina. Los métodos para ensayar la función efectora de los anticuerpos son bien conocidos en la técnica e incluyen aquellos que ensayan la CDC, la ADCC y la apoptosis. Véase, por ejemplo, Subbramanian et al. (2002) *J. Clin. Microbiol.* 40: 2141-2146; Ahman et al. (1994) *J. Immunol. Methods* 36: 243-254; Brezicka et al. (2000) *Cancer Immunol. Immunother.* 49: 235-242; Gazzano-Santoro et al. (1997) *J. Immunol. Methods* 202: 163-171; Prang et al. (2005) *British J. Cancer* 92: 342-349; Shan et al. (1998) *Blood* 92: 3756-3771; Ghetie et al. (2001) *Blood* 97: 1392-1398; y Mathas et al. (2000) *Cancer Research* 60:7170-7176.

En la técnica es conocido que el estatus de glicosilación de la porción Fc de una molécula de anticuerpo desempeña un papel clave en la determinación de si un anticuerpo tendrá función efectora. Véase, por ejemplo, Tao y Morrison (1987) *J. Immunol.* 143: 2595-601; Wright y Morrison (1997) *Trends in Biotech.* 15: 26-32; Wright y Morrison (1998) *J. Immunol.* 160: 3393-402; Mimura et al. (2000) *Mol. Immunol.* 37: 697-706; Jefferis y Lund (2002) *Immunol. Lett.* 82: 57-65; Krapp et al. (2003) *J. Mol. Biol.* 325: 979-89; y Jefferis (2005) *Biotechnol. Prog.* 21: 11-16. La glicosilación de anticuerpos producidos recombinantemente varía dependiendo del sistema de expresión usado. Véase, por ejemplo, Raju et al. (2000) *Glycobiology* 10: 477-86; Wright y Morrison (1997) *Trends in Biotech.* 15: 26-32. Además, cuando el modelo de *N*-glicosilación de un anticuerpo monoclonal producido en un mamífero es alterado para reducir o agotar el residuo de fucosa de núcleo α (1,6)-ligado, el anticuerpo monoclonal exhibe un aumento de la función efectora, particularmente un aumento de la actividad de ADCC. Véase, por ejemplo, la Patente de EE.UU. Nº: 6.946.292. Para algunos anticuerpos monoclonales producidos en mamíferos, cuando el modelo de *N*-glicosilación es alterado para reducir o agotar los residuos de β (1,4)-galactosa unidos a los brazos de 1,3 y/o 1,6 manosa, la activación de la citotoxicidad dependiente de complemento (CDC) contra las células diana que portan antígeno puede reducirse sin alterar otras actividades funcionales del anticuerpo, incluyendo la actividad de ADCC. Véase, por ejemplo, Boyd et al. (1995) *Mol. Immunol.* 32: 1311-1318.

Los anticuerpos que tienen actividad de reconocimiento de antígenos, y en algunas realizaciones una mejor función

efectora, pueden ser producidos por una planta superior hospedante, tal como una lenteja de agua, que ha sido transformada establemente de la manera establecida en la presente memoria para alterar su maquinaria de glicosilación. Por consiguiente, la presente invención proporciona métodos para producir un anticuerpo monoclonal recombinante, que incluye un anticuerpo monoclonal que tiene una función efectora mejorada, en donde el anticuerpo es producido recombinantemente dentro de una planta que tiene un modelo de N-glicosilación alterado de glicoproteínas endógenas y heterólogas producidas en la presente memoria de tal modo que dichas glicoproteínas exhiban una reducción de la cantidad de residuos de xilosa β 1,2-ligados y residuos de fucosa α 1,3-ligados específicos de planta unidos a los N-glicanos de las mismas. Cuando los anticuerpos tienen cantidades reducidas de residuos de fucosa α 1,3-ligados unidos a los N-glicanos de los mismos, los anticuerpos pueden tener una actividad ADCC incrementada respecto a los anticuerpos producidos en una planta de control que no ha sido modificada genéticamente para inhibir la expresión o la función de la FucT.

También se abarca la producción de anticuerpos monoclonales recombinantes que tienen función efectora, y en algunas realizaciones, una función efectora mejorada, en donde los anticuerpos son producidos en un sistema de expresión en lentejas de agua que ha sido modificado genéticamente para inhibir la expresión o la función de la FucT de la SEC ID N°: 3 y/o la XylT de SEC ID N°: 6, y cualquier isoforma de la misma, por ejemplo, la isoforma n°2 de XylT que comprende la secuencia establecida en la SEC ID N°: 21 (codificada por la SEC ID N°: 20). La planta que actúa como hospedante para la producción recombinante del anticuerpo monoclonal es un miembro de las *Lemnaceae* como se ha indicado en otra parte de la presente memoria, por ejemplo, una planta *Lemna*, que comprende, por ejemplo, una casete de expresión de ARNi de XylT y/o una casete de expresión de ARNi de FucT descritas anteriormente integradas establemente dentro de su genoma. De esta manera, la presente invención proporciona un método para producir un anticuerpo monoclonal recombinante que tiene un modelo de N-glicosilación que se asemeja más estrechamente al observado en un sistema de expresión de hospedante mamífero, y con una función efectora mejorada, en donde el método comprende expresar una o más cadenas del anticuerpo en una lenteja de agua, o en una célula o nódulo de lenteja de agua, que ha sido modificada genéticamente para alterar la maquinaria de glicosilación de tal modo que el anticuerpo monoclonal producido recombinantemente exhibe una reducción de la unión de los residuos de xilosa β 1,2-ligados y residuos de fucosa α 1,3-ligados específicos de planta con los N-glicanos de los mismos, y cultivar la lenteja de agua, o la célula o nódulo de lenteja de agua, en condiciones adecuadas para la expresión del anticuerpo monoclonal. Por tanto, la presente invención se refiere a las composiciones de anticuerpos en las que el anticuerpo comprende glicanos N-ligados que son predominantemente de estructura de glicano G0, es decir composiciones de anticuerpos, que tienen perfiles de glicosilación que son "sustancialmente homogéneos" o "sustancialmente uniformes" o que tienen una "homogeneidad sustancial" tal como se ha definido anteriormente en la presente memoria. Las composiciones sustancialmente homogéneas para las especies de glicano G0 tienen un perfil de glicosilación sustancialmente homogéneo, en donde al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos el 99% de la cantidad total de especies de N-glicano que aparecen en el perfil de glicosilación correspondiente a la composición está representado por especies de glicano G0, con una cantidad traza de especies de N-glicano precursoras que aparecen en el perfil de glicosilación, es decir, cualquier especie de N-glicano precursora dada que está presente en el perfil de glicosilación está presente en menos de un 5%, preferiblemente menos de un 4%, menos de un 3%, menos de un 2%, menos de un 1% e incluso menos de un 0,5% o incluso menos de un 0,1% de la cantidad total de especies de N-glicano que aparecen en el perfil. Para dicha composición, una especie de N-glicano precursora representativa que aparece en su perfil de glicosilación serían, por ejemplo, las especies de N-glicano precursoras descritas anteriormente, Man3GlcNAc2, MGn (GlcNAc1Man3GlcNAc2 en donde GlcNAc1 está unida al brazo de 1,3 manosa) y GnM (GlcNAc1Man3GlcNAc2 en donde GlcNAc1 está unida al brazo de 1,6 manosa), en donde cualquiera de estas especies de N-glicano precursoras, individualmente o en combinación, pueden estar presentes.

En un ejemplo de ese tipo, la composición de anticuerpos monoclonales tiene un perfil de glicosilación sustancialmente homogéneo, en donde el 95,8% de la cantidad total de especies de N-glicano que aparecen en el perfil de glicosilación correspondiente a la composición está representado por las especies de glicano G0 (GlcNAc₂Man₃GlcNAc₂), apareciendo las siguientes especies de N-glicano precursoras en el perfil de glicosilación: Man₃GlcNAc₂ (0,67%), GlcNAcMan₃GlcNAc₂ (1,6%), GalGlcNAc₂Man₃GlcNAc₂ (1,2%), Man₆GlcNAc₂ (0,21%), Man₇GlcNAc₂ (0,30%) y Man₈GlcNAc₂ (0,28%). Esto se puede comparar con la composición de anticuerpos monoclonales obtenida del sistema de expresión de lenteja de agua "natural" en donde se expresa el mismo anticuerpo monoclonal pero en donde la maquinaria de glicosilación de la planta de lenteja de agua no ha sido modificada genéticamente para inhibir la expresión de XylT y FucT. Dicha composición de anticuerpo monoclonal "natural" tiene un perfil de glicosilación más heterogéneo que se caracteriza por dos especies de N-glicanos predominantes, es decir, G0XF³ y G0X, con varias especies de N-glicano precursoras representadas en cantidades traza. En dicho ejemplo, la composición de anticuerpo monoclonal "natural" tiene un perfil de glicosilación con las siguientes especies de N-glicano representadas en él: G0 (GlcNAc₂Man₃GlcNAc₂) (8,4%); G0X (GlcNAc₂[Xyl]Man₃GlcNAc₂) (17,2%); G0XF³ (GlcNAc₂[Xyl]Man₃[Fuc]GlcNAc₂) (67,4%); Man₃GlcNAc₂ (0,26%); GlcNAcMan₃GlcNAc₂ (0,40%); (Xyl)Man₃(Fuc)GlcNAc₂ (0,76%); GlcNAc₂Man₃(Fuc) GlcNAc₂ (2,1%); GlcNAc(Xyl)Man₃(Fuc)GlcNAc₂ (1,4%); Man₆GlcNAc₂ (0,21%); Man₇GlcNAc₂ (0,63%); Gal(Fuc)GlcNAc₂(Xyl)Man₃(Fuc)GlcNAc₂ (0,26%); Man₈GlcNAc₂ (0,61 %) y Man₉GlcNAc₂ (0,40%).

De esta manera, la invención se refiere a composiciones de anticuerpos "sustancialmente homogéneas" o "sustancialmente uniformes" o a composiciones de anticuerpos que tienen una "homogeneidad sustancial" tal como

se ha definido anteriormente en la presente memoria, es decir, composiciones en las que al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos 99% del anticuerpo presente en la composición está representado por la glicoforma G0, en donde todos los sitios de glicosilación anticipados (por ejemplo, cada uno de los residuos de Asn-297 de los dominios C_H2 de las cadenas pesadas de un anticuerpo de tipo IgG) están ocupados por las especies de glicano G0, con una cantidad traza de glicoformas precursoras presentes en la composición. En una composición de este tipo, las glicoformas precursoras se seleccionan del grupo que consiste en un anticuerpo que tiene una región Fc en la que el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de la otra cadena pesada está sin glicosilar; un anticuerpo que tiene una región Fc en la que el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de la otra cadena pesada tiene el glicano precursor GnM o MGn unido a Asn 297; y un anticuerpo que tiene una región Fc en la que el sitio de glicosilación de Asn 297 de cada uno de los dominios C_H2 tiene una especie de glicano G0 unida, con una tercera especie de glicano G0 unida a un sitio de glicosilación adicional dentro de la estructura de mAb; en donde hay presente una cantidad traza de estas glicoformas precursoras, es decir, las glicoformas precursoras representan menos del 5%, menos del 4%, menos del 3%, menos del 2%, menos del 1%, o incluso menos del 0,5%, o menos del 0,1% de las glicoformas totales presentes en la composición de anticuerpos.

Las composiciones de anticuerpos sustancialmente homogéneas en las que al menos el 80%, al menos el 85%, al menos el 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, o al menos el 99% del anticuerpo presente en la composición está representado por la glicoforma G0, representan "anticuerpos de glicano optimizado". Por "anticuerpos de glicano optimizado" se entiende que los anticuerpos de la invención han sido modificados genéticamente en su modelo de glicosilación de tal modo que son sustancialmente homogéneos para la glicoforma G0, que da lugar a un anticuerpo que tiene una función efectora Fc mejorada. Por "función efectora Fc mejorada" se entiende que dichos anticuerpos tienen una actividad ADCC aumentada respecto a los anticuerpos de la misma secuencia (es decir, los anticuerpos tienen la misma secuencia de aminoácidos) que, como resultado del proceso de producción, tienen un perfil de glicosilación más heterogéneo. Así, por ejemplo, los anticuerpos producidos en sistemas de expresión de hospedante de mamífero, por ejemplo las células de CHO, en células hospedantes de insecto, en células de levadura, o en otros sistemas de expresión de hospedantes vegetales que no han sido alterados genéticamente para inhibir la expresión de XylT y FucT tienden a tener perfiles de glicosilación más heterogéneos, y por tanto una mezcla de glicoformas, que pueden efectuar una función efector global del producto de anticuerpo. La glicoforma G0 de estas composiciones de anticuerpos proporciona de forma ventajosa una composición de anticuerpos que tiene una actividad ADCC asociada a la ausencia de residuos de fucosa. En algunas realizaciones, la actividad de ADCC aumenta 25 veces, 50 veces, 75 veces, 100 veces, 150 veces, 200 veces, 250 veces, 300 veces, 400 veces, 500 veces o incluso 1000 veces respecto a anticuerpos de la misma secuencia que tienen un perfil de glicosilación heterogéneo (es decir, con múltiples glicoformas presentes como glicoformas principales en la composición de anticuerpos). Además, la glicoforma G0 carece de residuos Gal terminales presentes en anticuerpos que tienen la glicoforma G2. Como tal, estas composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoforma G0 presentan relaciones ADCC/CDC incrementadas. Adicionalmente, dichas composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoforma G0 presentan un aumento de la unión a la FcγRIII, por ejemplo, FcγRIIIa, en donde la afinidad de unión aumenta aproximadamente 20 veces, 30 veces, 40 veces, 50 veces, 75 veces, hasta 100 veces sobre la observada en composiciones de anticuerpos de la misma secuencia que tienen un perfil de glicosilación heterogéneo, y por tanto una mezcla de glicoformas. Para enfermedades oncológicas y autoinmunes, los anticuerpos terapéuticos que tienen una afinidad de unión incrementada para receptores de Fc, por ejemplo, FcγRIII, se han correlacionado fuertemente con una eficacia incrementada y una respuesta mejorada al tratamiento.

Algunas realizaciones de composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoforma G0 presentan una actividad de CDC alterada cuando se comparan con la observada en composiciones de anticuerpos de la misma secuencia que tienen un perfil de glicosilación heterogéneo, y por tanto una mezcla de glicoformas. Por ejemplo, las composiciones de anticuerpos sustancialmente homogéneas de la invención predominantemente con la glicoforma G0 pueden presentar una actividad de CDC disminuida cuando se comparan con la observada en composiciones de anticuerpos de la misma secuencia que tienen un perfil de glicosilación heterogéneo, y por tanto una mezcla de glicoformas. Por tanto, la presente invención se refiere a composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoforma G0 y actividad CDC que está disminuida como mucho en un 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% o incluso 100% cuando se comparan con composiciones de anticuerpos de igual secuencia que tienen un perfil de glicosilación heterogéneo. Estas composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoforma G0 y una actividad CDC disminuida se pueden caracterizar adicionalmente por tener una actividad de ADCC incrementada cuando se compara con composiciones de anticuerpos de la misma secuencia que tienen un perfil de glicosilación heterogéneo.

Sin pretender establecer ninguna teoría o mecanismo de acción, una composición de anticuerpos de glicoforma G0 sustancialmente homogénea que tienen una actividad de CDC disminuida, y una actividad de ADCC similar o incrementada de la manera descrita anteriormente, puede proporcionar de manera ventajosa una citotoxicidad incrementada contra células diana a la vez que se reducen los potenciales efectos secundarios adversos que puedan estar asociados a la activación de complemento después de su administración. Reduciendo el potencial de

dichos efectos secundarios adversos, las composiciones de anticuerpos sustancialmente homogéneas que tienen predominantemente la glicoformas G0 pueden administrarse de manera ventajosa con mayores velocidades de infusión, reduciendo con ello el tiempo de dosificación de cualquier administración dada, y/o se pueden dosificar a mayores concentraciones iniciales si están garantizadas, con una preocupación reducida para activar efectos secundarios adversos asociados a la activación de complemento.

Por ejemplo, la activación de complemento desempeña un papel central en la patogénesis de efectos secundarios de primera dosis entre moderados y graves del tratamiento con el anticuerpo monoclonal anti-CD20 quimérico IDEC-C2B8 (IDEC Pharmaceuticals Corp., San Diego, California; disponible comercialmente bajo el nombre comercial Rituxan®, también conocido como rituximab). Véase, por ejemplo, van der Kolk et al. (2001) British J. Haematol. 115: 807-811. El anticuerpo rituximab del producto Rituxan® se expresa en células de ovario de hámster chino (CHO), y de este modo la composición de anticuerpos comprende un perfil de glicosilación heterogéneo (es decir, una mezcla de glicoformas). Se ha demostrado que la actividad de CDC del rituximab se correlaciona con el contenido de galactosa. De esta manera, según aumenta el número de residuos de galactosa desde 0-2 moles/mol de cadena pesada, el nivel de actividad de CDC aumenta desde el 80% (β -galactosidasa tratada para eliminar todos los residuos de β (1,4)-galactosa de los brazos de 1,3 y 1,6 manosa de los *N*-glicanos unidos a Asn 297 de los dominios C_H2 de las cadenas pesadas) hasta el 150% (UDP galactosil transferasa tratada para asegurar que los residuos de β (1,4)-galactosa estén unidos a ambos brazos de 1,3 y 1,6 manosa de los *N*-glicanos unidos a sitios de Asn 297) del máximo observado para el anticuerpo que tiene 1 mol de galactosa/mol de cadena pesada (véase, IDEC BLA 97-0260 en la página web fda.gov/Cder/biologics/review/ritugen112697, disponible en internet). Una composición de anticuerpos anti-CD20 sustancialmente homogénea que comprende un anticuerpo anti-CD20 que tiene la misma secuencia que el rituximab y que tiene predominantemente la glicoforma G0 tendría de forma ventajosa una actividad de CDC disminuida, reduciendo con ello el potencial de efectos secundarios adversos normalmente asociados a la activación de complemento tras la administración de anticuerpos cuando la composición de anticuerpos comprende un perfil de glicosilación heterogéneo (es decir, una mezcla de glicoformas).

Por tanto, una composición de anticuerpos de glicoforma G0 sustancialmente homogénea que tiene una actividad de CDC disminuida, y una actividad de ADCC igual o incrementada, puede usarse de manera ventajosa en aplicaciones terapéuticas que hasta el momento no han sido adecuadas, aconsejables o eficaces para una o más poblaciones de pacientes como resultado de complicaciones debidas a efectos secundarios adversos asociados normalmente a la activación de complemento tras la administración de una composición de anticuerpos de la misma secuencia que comprende un perfil de glicosilación heterogéneo (es decir, una mezcla de glicoformas). Dichos efectos secundarios incluyen, aunque sin limitación, efectos secundarios de moderados a severos que están asociados a una administración de un anticuerpo por primera vez y/o rápida, que incluyen, por ejemplo, fiebre y/o temblores, náusea, disnea, sarpullidos, y otros similares. Véase, por ejemplo, van der Kolk et al. (2001) British J. Haematol. 115: 807-811 y las referencias citadas en la misma; Winkler et al. (1999) Blood 94: 2217-2224). De esta manera, la presente invención proporciona un método para reducir uno o más efectos secundarios adversos relacionados con la activación de complemento tras la administración de un anticuerpo, por ejemplo, un anticuerpo monoclonal, comprendiendo dicho método la administración del anticuerpo como una composición de anticuerpos sustancialmente homogénea como se ha definido anteriormente en la presente memoria, y por tanto al menos un 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o al menos un 99% del anticuerpo presente en la composición está representado por la glicoforma G0, estando presentes en la composición las glicoformas precursoras en cantidades traza. En algunas realizaciones, al menos el 90%, al menos el 95% o al menos el 99% de los anticuerpos presentes en la composición están representados por la glicoformas G0, estando presentes en la composición las glicoformas precursoras en cantidades traza.

Como resultado de su función efectora Fc incrementada, las composiciones de anticuerpos sustancialmente homogéneas que tiene predominantemente la glicoforma G0 proporcionan las oportunidades para rutas de administración nuevas y mejoradas, por ejemplo, extendiendo las posibles rutas de administración correspondientes a anticuerpo terapéuticos conocidos a otras rutas de administración más allá de la infusión y la administración intravenosa, por ejemplo, a administración subcutánea. Además, como resultado de su potencia incrementada, las composiciones de anticuerpos se pueden dosificar a concentraciones menores, o se puede dosificar a volúmenes más pequeños, y con menos frecuencia. Una reducción del volumen de composición de anticuerpos administrada es particularmente ventajosa en los casos en los que eventos adversos producidos por reacciones a la infusión con un anticuerpo monoclonal están relacionados con el volumen. La potencia incrementada de las composiciones de anticuerpos de la invención también abre las puertas a nuevas indicaciones clínicas para las dianas de mAb existentes que no han respondido a la terapia con anticuerpos con composiciones de anticuerpos de glicoforma más heterogéneas, y proporciona la capacidad para volver a considerar dianas de mAb no desarrolladas previamente.

Los anticuerpos monoclonales producidos de acuerdo con los métodos de la presente invención pueden estar contenidos en una composición que comprende un vehículo farmacéuticamente aceptable. Dichas composiciones son útiles en un método de tratamiento de un sujeto que necesite un anticuerpo que tenga una función efectora, y en algunas realizaciones, una función efectora mejorada en donde la expresión de FucT ha sido la diana de la inhibición. De esta manera, los anticuerpos monoclonales producidos en una planta, por ejemplo, una lenteja de agua, transformada establemente de acuerdo con los métodos de la presente invención se pueden administrar a un sujeto que lo necesite.

Dentro de un hospedante vegetal de lenteja de agua modificada genéticamente se puede producir cualquier anticuerpo de acuerdo a los métodos de la invención. En una realización, el anticuerpo es un anticuerpo terapéutico. Se han aprobado anticuerpos para el tratamiento de una serie de trastornos, y muchas terapias adicionales con anticuerpos se encuentran en desarrollo. Véase, Brekke y Sandlie (2003) Nat. Rev. Drug. Discov. 2(3): 240. Los ejemplos de anticuerpos terapéuticos que pueden expresarse mediante los métodos de la invención incluyen, aunque sin limitación, los anticuerpos anti-CD3 dirigidos contra el antígeno CD3 (por ejemplo, OKT®3); los anticuerpos anti-plaquetas gpIIb/IIIa (por ejemplo, abciximab (previamente conocido como Fab c7E3), distribuido como ReoPro®; los anticuerpos anti-EpCAM (por ejemplo, Panorex®); los anticuerpos anti-CD20 dirigidos contra el antígeno CD20 (por ejemplo, rituximab, distribuido como Rituxan®; véase la Patente de EE.UU. N° 5.736.137); el receptor anti-IL-2 (por ejemplo, Zenapax®, Simulect®); anti-ERBB2 (por ejemplo, trastuzumab, distribuido como Herceptin®; véase la Patente de EE.UU. N° 6.165.464); anti-TNF- α (por ejemplo, infliximab, distribuido como Remicade®); anti-proteína-F (por ejemplo, Synagis®); anticuerpos anti-CD30 dirigidos contra el antígeno CD30 (véase, por ejemplo, el anticuerpo 5F11 descrito en Borchmann et al. (2003) Blood 102: 3737-3742 y en el Ejemplo 6 mostrado en la presente memoria más adelante, véase también el documento WO 03/059282 y la Publicación de Solicitud de Patente de EE.UU. N° 2004/0006215; el anticuerpo SGN-30 descrito en Wahl et al. (2002) Cancer Res. 62: 3736-3742, una versión quimérica del anticuerpo anti-CD30 AC10, véase también la Publicación de Solicitud de Patente de EE.UU. N° 20040018194 y la Patente de EE.UU. N° 7.090.843); los anticuerpos anti-CD33 dirigidos contra el antígeno CD33 (por ejemplo, Mylotarg®; véase la Patente de EE.UU. N° 5.733.001); los anticuerpos anti-CD52 dirigidos contra el antígeno CD52 (por ejemplo, alemtuzumab, distribuido como Campath®; véase la Patente de EE.UU. N°: 5.846.534); anti-CD20 (por ejemplo, Zevalin®); anti-IgE Fc (por ejemplo, omalizumab, comercializado como Xolair®); anti-VEGF (por ejemplo, bevacizumab, distribuido como Avastin®; véase, Presta et al. (1997) Cancer Res. 57: 4593-9); anti-EGF-R (por ejemplo, cetuximab, distribuido como Erbitux®); anti-CD11a (por ejemplo, efalizumab, distribuido como Raptiva®), anti-IL-8; anti-C5; anti-TNF- α (por ejemplo, adalimumab, distribuido como Humira®); anti-TGF- β 2; receptor anti-IL2 (por ejemplo, daclizumab (*Zenapax*) y basiliximab (*Simulect*)); anti- α -4-integrina; anti-CD4; anti-CD2; anti-CD19 (por ejemplo, MT103, un anticuerpo biespecífico); anticuerpo anti-CD22 dirigido contra el antígeno CD22 (por ejemplo, el anticuerpo monoclonal BL-22); anticuerpo anti-CD23 dirigido contra el antígeno CD23 de células B malignas (por ejemplo, IDEC-152); anticuerpo anti-CD80 dirigido contra el antígeno CD80 (por ejemplo, IDEC-114); anticuerpo anti-CD38 dirigido contra el antígeno CD38 de células B malignas; anticuerpo α -M-CSF dirigido contra factor estimulante de colonia de macrófagos; anticuerpos dirigidos contra el activador de receptor de factor-kappaB nuclear (RANK) y su ligando (RANKL), que están sobreexpresados en el mieloma múltiple; anticuerpos anti-CD40 (por ejemplo, SGN-40) dirigidos contra el antígeno CD40 de células B malignas; anticuerpos dirigidos contra el ligando receptor 1 inductor de apoptosis relacionada con factor de necrosis tumoral (TRAIL-R1) (por ejemplo, el anticuerpo monoclonal humano agonístico HGS-ETR1) y TRAIL-R2 expresado en una serie de tumores sólidos y tumores de origen hematopoiético; anticuerpos anti-HBV; o anticuerpos anti-MHC de clase II. Muchos de estos anticuerpos son conocidos y requieren de actividad efectora para su función, CDC y/o ADCC, por ejemplo, Herceptin®, Rituxan®, Mylotarg® y Campath®.

En algunas realizaciones, la planta hospedante que ha sido modificada genéticamente para alterar su maquinaria de glicosilación de la manera establecida en la presente memoria, por ejemplo, una lenteja de agua, puede expresar α -2b-interferón biológicamente activo, por ejemplo precursor de α -2b-interferón (N° de acceso de proteína NCBI AAB9402) o variantes biológicamente activas del mismo. Los ejemplos de variantes biológicamente activas del α -2b-interferón humano se conocen en la técnica. Véase, por ejemplo, el documento WO 2005/035767, que describe versiones truncadas de α -2b-interferón; la Patente Europea EP 211148B1; y las Patentes de EE.UU. N° 4.748.233, 4.801.685, 4.816.566, 4.973.479, 4.975.276, 5.089.400, 5.098.703, 5.231.176 y 5.869.293. En otras realizaciones, la planta hospedante, por ejemplo, una lenteja de agua, puede expresar hormona del crecimiento humano madura biológicamente activa con o sin su péptido señal acompañante. La planta hospedante también puede expresar variantes biológicamente activas de la hormona de crecimiento humana. Véase, por ejemplo, el documento WO 02/10414, dirigido a la expresión de polipéptidos biológicamente activos usando un sistema de expresión de lenteja de agua.

En otras realizaciones adicionales, la planta hospedante, por ejemplo, una lenteja de agua, puede expresar plasminógeno, microplasminógeno, o variantes biológicamente activas del mismo. El plasminógeno o el microplasminógeno que van a ser expresados en la planta hospedante pueden proceder de cualquier fuente de mamífero. En algunas realizaciones, el plasminógeno o el microplasminógeno son humanos o porcinos. Véase, por ejemplo, el documento WO 2005/078109, dirigido a la expresión de dichas proteínas usando un sistema de expresión de lenteja de agua.

Por "variante biológicamente activa" de dichos polipéptidos heterólogos se entiende un polipéptido derivado del polipéptido nativo por eliminación (el denominado truncamiento) o adición de uno o más aminoácidos al extremo N-terminal y/o C-terminal de la proteína nativa; eliminación o adición de uno o más aminoácidos en uno o más sitios de la proteína nativa; o sustitución de uno o más aminoácidos en uno o más sitios de la proteína nativa. Las variantes de un polipéptido heterólogo de interés abarcadas por la presente invención mantienen la actividad biológica de los polipéptidos nativos.

Por ejemplo, las variantes biológicamente activas de α -2b-interferón continúan poseyendo la actividad biológica deseada del α -2b-interferón nativo, incluyendo la capacidad para aumentar la resistencia frente a infección vírica, o la capacidad para modular la transcripción de dianas génicas reguladas por α -2b-interferón. Dichas variantes

biológicamente activas pueden ser el resultado de, por ejemplo, polimorfismo genético o de manipulación humana. Las variantes biológicamente activas de una proteína heteróloga de interés tendrán al menos aproximadamente un 50%, 60%, 65%, 70%, generalmente al menos aproximadamente un 75%, 80%, 85%, preferiblemente al menos aproximadamente un 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, y más preferiblemente al menos aproximadamente un 98%, 99% o más de identidad de secuencia con respecto a la secuencia de aminoácidos correspondiente a la proteína heteróloga de interés. Así, por ejemplo, una variante biológicamente activa del α -2b-interferón tendrá al menos aproximadamente el 50%, 60%, 65%, 70%, generalmente al menos aproximadamente el 75%, 80%, 85%, preferiblemente al menos aproximadamente 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, y más preferiblemente al menos aproximadamente 98%, 99% o más de identidad de secuencia con respecto a la secuencia de aminoácidos correspondiente al precursor de α -2b-interferón (N° de acceso de proteína NCBI AAB59402) o de α -2b-interferón humano maduro (aminoácidos 24-188 del número de acceso de proteína NCBI AAB9402), en donde la identidad de secuencia se determina usando los parámetros identificados anteriormente en la presente memoria. Por tanto, una variante biológicamente activa de una proteína heteróloga nativa de interés puede diferir de dicha proteína en sol 1-15 residuos de aminoácidos, solo 1-10, solo 6-10, solo 5, solo 4, 3, 2 o incluso 1 residuo de aminoácido.

En una realización, la planta hospedante que ha sido modificada genéticamente para alterar su maquinaria de glicosilación de la manera establecida en la presente memoria es una planta de lenteja de agua, ó célula o nódulo de lenteja de agua, transformada establemente que expresa polipéptidos biológicamente activos que no se pueden producir comercialmente de manera efectiva mediante los sistemas de expresión génica existentes, debido a limitaciones económicas o logísticas, o ambas. Por ejemplo, algunas proteínas no pueden expresarse en sistemas de mamíferos debido a que la proteína interfiere con la viabilidad celular, la proliferación celular, la diferenciación celular o el ensamblaje de proteínas en células de mamíferos. Dichas proteínas incluyen, aunque sin limitación, la proteína de retinoblastoma, p53, angiostatina y leptina. La presente invención puede emplearse de manera ventajosa para producir proteínas reguladoras de mamíferos; es improbable, dada la gran distancia evolucionaria entre las plantas superiores y los mamíferos, que estas proteínas interfieran con procesos reguladores de la lenteja de agua. La lenteja de agua transgénica también puede usarse para producir grandes cantidades de proteínas tales como albúmina de suero (en particular, albúmina de suero humano), hemoglobina y colágeno, que comprometen las capacidades de producción de los sistemas de expresión existentes.

Finalmente, los sistemas de plantas superiores pueden modificarse para producir proteínas multiméricas biológicamente activas (por ejemplo, anticuerpos monoclonales, hemoglobina, oxidasa P450 y colágeno, y similares) mucho más fácilmente que los sistemas de mamíferos. Un ejemplo de estrategia para producir proteínas multiméricas biológicamente activas en lentejas de agua usa un vector de expresión que contiene los genes que codifican todas las subunidades de polipéptido. Véase, por ejemplo, Daring y col. (1990) Plant Mol. Biol. 15: 281 y van Engelen y col. (1994) Plant Mol. Biol. 26: 1701. La casete de expresión que comprende el polinucleótido inhibidor de XylT y/o FucT puede introducirse en dicho vector. A continuación, dicho vector se introduce en células de lenteja de agua usando cualquier método de transformación conocido, tal como el bombardeo balístico o la transformación mediada por *Agrobacterium*. Este método da como resultado líneas celulares clonales que expresan todos los polipéptidos necesarios para ensamblar la proteína multimérica, así como las secuencias inhibitoras de XylT y/o FucT que alteran el modelo de glicosilación de los *N*-glicanos de las glicoproteínas. Por consiguiente, en algunas realizaciones, la lenteja de agua transformada contiene uno o más vectores de expresión que codifican una cadena pesada y ligera de un anticuerpo monoclonal o fragmento de anticuerpo Fab', y una casete de expresión que comprende el polinucleótido inhibidor de XylT y/o FucT, y el anticuerpo monoclonal o fragmento de anticuerpo se ensambla en una planta de lenteja de agua a partir de la cadena pesada y ligera expresada.

Una variación de esta estrategia es preparar construcciones de gen individual, mezclar el ADN de dichas construcciones, y a continuación suministrar la mezcla de ADNs a células vegetales usando bombardeo balístico o transformación mediada por *Agrobacterium*. Como variación adicional, algunos o todos los vectores pueden codificar más de una subunidad de la proteína multimérica (es decir, de tal modo que se tenga que cruzar un número de clones de plantas de lenteja de agua menor que el número de subunidades de la proteína multimérica). En una realización alternativa, cada clon de lenteja de agua ha sido modificado genéticamente para alterar su maquinaria de glicosilación y expresa al menos una de las subunidades de la proteína multimérica, y los clones de lenteja de agua que secretan cada subunidad se cultivan juntos y la proteína multimérica se ensambla en el medio a partir de las diversas subunidades secretadas. En algunos casos, puede ser deseable producir menos que todas las subunidades de una proteína multimérica, o incluso una subunidad de proteína individual, en una planta de lenteja de agua transformada o un cultivo de nódulo de lenteja de agua, por ejemplo para procesos industriales o químicos, o con fines diagnósticos, terapéuticos o de vacunación.

En algunas realizaciones de la invención, la planta transgénica hospedante de interés es un "alto expresador" de una glicoproteína descrita en la presente memoria, que incluye, por ejemplo, las glicoproteínas que comprenden glicanos *N*-ligados que predominantemente son de estructura de glicano G0. Por "alto expresador" se entiende la planta transgénica hospedante que ha sido modificada genéticamente para producir las glicoproteínas descritas en la presente memoria es capaz de producir la glicoproteína de interés a un nivel tal que la glicoproteína de interés representa al menos el 5% o más de la proteína soluble total producida en la planta transgénica hospedante. En algunas realizaciones, un "alto expresador" es una planta transgénica hospedante que ha sido modificada para producir las glicoproteínas descritas en la presente memoria de tal modo que la glicoproteína de interés representa al menos el 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15% o más de la proteína soluble total producida en

la planta transgénica hospedante. Por tanto, por ejemplo, en una realización, la planta transgénica hospedante es una lenteja de agua que ha sido modificada para inhibir la expresión de XylT y FucT, y la lenteja de agua transgénica es un alto expresador de una glicoproteína descrita en la presente memoria. En algunas de estas realizaciones, la lenteja de agua transgénica es un alto expresador de un anticuerpo monoclonal de glicano optimizado que tienen la glicoforma G0 predominante descrita anteriormente en la presente memoria. En otras realizaciones adicionales, la lenteja de agua transgénica expresa el anticuerpo monoclonal de glicano optimizado de tal modo que la glicoproteína representa aproximadamente el 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15% ó más, de la proteína soluble total.

Humanización adicional de Glicoproteínas:

En algunas realizaciones, puede ser deseable que las glicoproteínas comprendan *N*-glicanos complejos que tienen residuos de β (1,4)-galactosa terminales unidos a los brazos de 1,3 y/o 1,6 manosa. Sin pretender establecer una teoría, estos residuos terminales de galactosa pueden contribuir a la función terapéutica y/o a la actividad farmacocinética de una glicoproteína. Se reconoce que los métodos de la presente invención se pueden igualar a otros métodos conocidos en la técnica para modificar adicionalmente las glicoproteínas de la invención de tal modo que uno o más de los *N*-glicanos ligados comprenden uno o más residuos terminales de galactosa, es decir, en donde uno o más de los *N*-glicanos está representado por especies de glicano G1 o G2. De esta manera, las composiciones de glicoproteína, que incluyen los anticuerpos monoclonales descritos en la presente memoria, se pueden modificar, por ejemplo, enzimáticamente con el uso de una enzima glicosiltransferasa para obtener glicoproteínas que tienen un perfil de glicosilación sustancialmente homogéneo para las especies de glicano G1, preferiblemente para la G2. Véase, por ejemplo, la Publicación de Solicitud de Patente de EE.UU. N° 2004/0191256, que muestra la modificación de galactosiltransferasa de una glicoproteína sustrato para obtener una glicoproteína en donde sustancialmente todas las especies de glicano *N*-ligadas son de la forma G2. De esta manera, la glicoproteína de interés se puede hacer reaccionar con una galactosa activada en presencia de una galactosiltransferasa y una sal metálica. La galactosiltransferasa puede ser una β 1,4-galactosiltransferasa (GalT) de mamífero, por ejemplo, GalT humana, y la galactosa activada puede ser, por ejemplo, UDP-galactosa.

Alternativamente, las plantas transgénicas de la invención que tienen expresión de FucT y XylT silenciada de la manera establecida en la presente memoria se puede modificar adicionalmente en su maquinaria de glicosilación de tal modo que exprese una galactosiltransferasa y una eficientemente el residuo terminal de galactosa a los *N*-glicanos de las glicoproteínas endógenas y heterólogas producidas en la presente memoria. De esta manera, las plantas transgénicas de la invención pueden modificarse adicionalmente introduciendo una construcción de nucleótido, por ejemplo, una casete de expresión, que proporcione la expresión de una galactosiltransferasa. La galactosiltransferasa puede ser una β 1,4-galactosiltransferasa (GalT) de mamífero, por ejemplo, GalT humana (véase, por ejemplo, la Patente de EE.UU. N°: 6.998.267) o una GalT híbrida (véase, por ejemplo, el documento WO 03/078637) que comprende al menos una porción de una región de cola-transmembrana-tallo citoplásmica de una primera glicosiltransferasa (por ejemplo, una glicosiltransferasa vegetal tal como la xilosiltransferasa, la *N*-acetilglicosaminiltransferasa o la fucosiltransferasa) y al menos una porción de una región catalítica de una segunda glicosiltransferasa (por ejemplo, glicosiltransferasa de mamífero, por ejemplo, GalT humana). Silenciando la expresión de XylT y FucT en una planta, por ejemplo, una lenteja de agua, y proporcionando la expresión de GalT, por ejemplo, GalT humana, o una enzima híbrida que comprende una porción del dominio catalítico de GalT, por ejemplo, GalT humana, en dicha planta, por ejemplo, lenteja de agua, es posible obtener plantas transgénicas que producen glicoproteínas, tanto endógenas como heterólogas, que tienen un modelo de glicosilación alterado, en donde los glicanos *N*-ligados unidos tienen una reducción en la unión de residuos de xilosa específicos de planta y de residuos de fucosa específicos de planta, y que comprende los residuos terminales de galactosa (es decir, especies de glicano G2). De esta manera, las glicoproteínas que tienen un perfil sustancialmente homogéneo para la especie de glicano G2, y/o que son sustancialmente homogéneas respecto a la glicoforma G2 se pueden obtener a partir de las plantas transgénicas de la invención.

En otras realizaciones, puede ser deseable modificar el modelo de glicosilación de las glicoproteínas, en donde los glicanos *N*-ligados unidos comprenden además un residuo de ácido siálico terminal unido a uno o ambos residuos de galactosa unidos a los brazos de 1,3 y 1,6 manosa. La adición de el(los) residuo(s) de ácido siálico puede ser un requisito para una estabilidad sostenida, y en algunos casos la función, de algunas proteínas terapéuticas.

Dependiendo del sistema de planta transgénica, se puede producir una sialización natural de las glicoproteínas. Así, se han publicado trabajos en la bibliografía que indican que células cultivadas de *Arabidopsis*, tabaco y *Medicago* sintetizan glicoproteínas sialiladas (Shah et al. (2003) Nat. Biotech. 21(12): 1470-1471; Joshi y Lopez (2005) Curr. Opin. Plant Biol. 8(2): 223-226). Más recientemente, se ha publicado que el arroz japonés expresa proteínas de tipo sialiltransferasa activas (Takashima et al. (2006) J. Biochem. (Tokyo) 139(2): 279-287). Por lo tanto, ahora existen trabajos ortogonales que indican que las plantas tienen la maquinaria requerida para sializar glicoproteínas.

Cuando se desea una modificación adicional del modelo de glicosilación de las glicoproteínas, en donde los glicanos *N*-ligados unidos a las mismas comprenden además un residuo de ácido siálico terminal unido a uno o ambos residuos de galactosa unidos a los brazos de 1,3 y 1,6 manosa, las plantas transgénicas de la invención pueden modificarse para expresar una β -1,4-galactosiltransferasa, por ejemplo, una β -1,4-galactosiltransferasa humana, y para expresar o sobreexpresar una sialiltransferasa. Así, por ejemplo, las plantas transgénicas pueden modificarse

adicionalmente para expresar una sialiltransferasa tal como α -2,3 y/o α -2,6 sialiltransferasa. Véase, por ejemplo, el documento WO 2004/071177; y Wee et al. (1998) Plant Cell 10: 1759-1768. Alternativamente, las plantas transgénicas de la invención pueden modificarse para expresar una β -1,4 galactosiltransferasa, por ejemplo, β -1,4 galactosiltransferasa humana, y para expresar cualquier otra enzima que sea deficiente en el mecanismo de ácido siálico de la planta hospedante. La(s) estrategia(s) empleada(s) puede(n) determinarse después de una investigación inicial sobre si la planta hospedante particular, por ejemplo una lenteja de agua, expresa de forma natural *N*-glicanos que contienen ácido siálico en glicoproteínas nativas o producidos recombinantemente. Por ejemplo, si no hay evidencia de la presencia de los residuos de ácido siálico terminales en los *N*-glicanos de las glicoproteínas producidas dentro de la planta hospedante transgénica, particularmente una planta hospedante transgénica modificada para expresar una β -1,4 galactosiltransferasa, entonces se podrían emplear una o ambas de estas estrategias para lograr una sialización terminal de los *N*-glicanos de las glicoproteínas producidas dentro de la planta hospedante transgénica de interés.

También se reconoce que para algunas glicoproteínas producidas en las plantas transgénicas de la invención, puede ser deseable tener el residuo de α 1-6-fucosa de mamífero unido a la estructura de núcleo de trimanosa ($\text{Man}_3\text{GlcNAc}_2$) de las especies de *N*-glicano unidas a la misma. En dichas realizaciones, las plantas transgénicas de la invención pueden modificarse genéticamente adicionalmente para expresar una α 1-6 fucosiltransferasa, por ejemplo, α 1-6 fucosiltransferasa humana, usando métodos de glicoingeniería conocidos en la técnica.

Las glicoproteínas producidas según los métodos de la presente invención se pueden recolectar a partir de las células hospedantes en las que se han producido recombinantemente con el objetivo de obtenerlas en su forma aislada o pura. De esta manera, las glicoproteínas producidas recombinantemente son aisladas de las células hospedantes usando cualquier medio convencional conocido en la técnica y se purifican, por ejemplo, mediante cromatografía, electroforesis, diálisis, extracción disolvente-disolvente, y otros similares. Así, la presente invención también proporciona glicoproteínas purificadas, que incluyen composiciones de anticuerpos monoclonales, en donde las glicoproteínas tienen perfiles de glicosilación sustancialmente homogéneos, y son sustancialmente homogéneas respecto a la glicofórmula G0. Estas glicoproteínas purificadas están sustancialmente libres de material celular hospedante, e incluyen preparaciones de glicoproteína que tienen menos de aproximadamente el 30%, 20%, 10%, 5% ó 1% (en base a peso seco) de proteína contaminante, tal como se ha indicado anteriormente en la presente memoria. En algunas realizaciones, estas glicoproteínas purificadas pueden incluir al menos un 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta 30% (en base a peso seco) de proteína contaminante. Además, para las glicoproteínas purificadas producidas recombinantemente, el medio de cultivo representa de manera óptima menos de aproximadamente el 30%, 20%, 10%, 5% ó 1% (en base a peso seco) de precursores químicos o productos químicos sin interés proteínico en la preparación de glicoproteína purificada, tal como se ha indicado anteriormente en la presente memoria. Por tanto, en algunas realizaciones, los componentes de medio de cultivo en estas glicoproteínas purificadas pueden representar al menos el 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta aproximadamente 30% (en base a peso seco) de precursores químicos o productos químicos sin interés proteínico en la preparación de glicoproteína purificada. En algunas realizaciones, el aislamiento y la purificación dan como resultado la recuperación de glicoproteína purificada que está libre de proteína hospedante contaminante, libre de componentes del medio de cultivo y/o libre de ambos, proteína hospedante contaminante y componentes del medio de cultivo.

Por tanto, el sistema hospedante de expresión de proteína es una lenteja de agua, y la glicoproteína purificada obtenida del hospedante vegetal está sustancialmente libre de material celular de la planta, que incluye realizaciones en las que las preparaciones de glicoproteínas tienen menos de aproximadamente 30%, 20%, 10%, 5% ó 1% (en base a peso seco) de proteína vegetal contaminante. En otras realizaciones, el medio de cultivo de la planta representa menos de aproximadamente el 30%, 20%, 10%, 5% ó 1% (en base a peso seco) de precursores químicos o productos químicos sin interés proteínico en la glicoproteína purificada.

En algunas realizaciones, dichas glicoproteínas purificadas obtenidas a partir de un hospedante de lenteja de agua pueden incluir al menos 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta 30% (en base a peso seco) de proteína vegetal contaminante. En otras realizaciones, los componentes del medio de cultivo de planta en dichas glicoproteínas purificadas pueden representar al menos 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta aproximadamente 30% (en base a peso seco) de precursores químicos o productos químicos sin interés proteínico en la glicoproteína purificada. En algunas realizaciones, el aislamiento y la purificación a partir de la planta hospedante dan como resultado la recuperación de glicoproteína purificada que está libre de proteína vegetal contaminante, libre de componentes del medio de cultivo de plantas y/o libre de ambos, proteína vegetal contaminante y componentes del medio de cultivo de plantas.

Casetes de Expresión:

Según la presente invención, las plantas superiores transformadas establemente, por ejemplo, una lenteja de agua transformada establemente, se obtienen mediante la transformación con un polinucleótido de interés contenido en

una casete de expresión. Dependiendo del objetivo, el polinucleótido de interés puede ser uno que codifique un polipéptido de FucT o XylT de interés, por ejemplo, que codifique el polipéptido establecido en la SEC ID N°: 3 (FucT) o la SEC ID N°: 6 ó 21 (XylT), o una variante del mismo, proporcionando de este modo la expresión de dichos polipéptidos en una célula, por ejemplo una célula vegetal, o puede ser un polinucleótido inhibidor de FucT o XylT que es capaz de inhibir la expresión o la función del polipéptido de FucT o XylT cuando se introduce de manera estable en una célula, por ejemplo en una célula vegetal de interés.

Por tanto, en algunas realizaciones, los polinucleótidos de FucT y/o XylT, que incluyen los establecidos en las SEC ID N°: 1 y 2 (FucT) y las SEC ID N°: 4, 5, 19 y 20 (XylT) y los fragmentos y variantes de los mismos, se usan para construir casetes de expresión que comprenden un polinucleótido inhibidor de FucT y/o XylT tal como se ha definido anteriormente en la presente memoria. La introducción de manera estable de dicha casete de expresión en una planta o célula vegetal de interés puede proporcionar la inhibición de la expresión o la función de los polipéptidos de FucT y/o XylT, que incluyen los establecidos en la SEC ID N°: 3 (FucT) y las SEC ID N°: 6 ó 21 (XylT) y las variantes de los mismos, alterando con ello el modelo de glicosilación de *N*-glicano de las glicoproteínas endógenas y heterólogas de la planta o célula vegetal transformadas establemente con la casete de expresión.

En algunas realizaciones, la planta o célula vegetal que es transformada establemente con una casete de expresión que comprende un polinucleótido inhibidor de FucT y/o XylT también ha sido transformado establemente con una casete de expresión que proporciona la expresión de un polipéptido heterólogo de interés, por ejemplo, una proteína de mamífero de interés, que incluye las proteínas de mamífero de interés farmacéutico tal como se ha indicado anteriormente en la presente memoria. La casete de expresión que proporciona la expresión de un polipéptido heterólogo de interés se puede proporcionar sobre el mismo polinucleótido (por ejemplo, sobre el mismo vector de transformación) para la introducción en una planta, o sobre un polinucleótido diferente (por ejemplo, sobre diferentes vectores de transformación) para la introducción en la planta o célula vegetal de interés al mismo tiempo o a tiempos diferentes, mediante el mismo método de introducción o con métodos diferentes, por ejemplo, mediante el mismo o diferente método de transformación.

Las casetes de expresión comprenden elementos de control de la expresión que al menos comprende una región de inicio transcripcional (por ejemplo, un promotor) ligado al polinucleótido de interés, es decir, un polinucleótido que codifica un polipéptido de FucT o XylT de la invención, un polinucleótido inhibidor de FucT y/o XylT, o un polinucleótido que codifica un polipéptido heterólogo de interés, por ejemplo, una proteína de mamífero. Dicha casete de expresión se proporciona con una pluralidad de sitios de restricción para la inserción del polinucleótido o polinucleótidos de interés (por ejemplo, un polinucleótido de interés, dos polinucleótidos de interés, etc.) que van a estar bajo regulación transcripcional del promotor y otros elementos de control de la expresión. En realizaciones particulares, el polinucleótido que va a ser transferido contiene dos o más casetes de expresión; cada una de las cuales codifica al menos un polinucleótido de interés.

Por "elemento de control de la expresión" se entiende una región reguladora de ADN, que normalmente comprende una caja TATA, capaz de dirigir ARN polimerasa II, o en algunas realizaciones, ARN polimerasa III, para iniciar la síntesis de ARN en el sitio de iniciación de la transcripción apropiado para una secuencia codificadora particular. Un elemento de control de la expresión puede comprender adicionalmente otras secuencias de reconocimiento generalmente posicionadas por encima o a 5' de la caja TATA, que influyen (por ejemplo, potencian) en la velocidad de iniciación de la transcripción. Además, un elemento de control de la expresión puede comprender adicionalmente secuencias posicionadas generalmente por debajo o a 3' de la caja TATA, que influyen (por ejemplo, potencian) en la velocidad de iniciación de la transcripción.

La región de iniciación transcripcional (por ejemplo, un promotor) puede ser nativa u homóloga o extraña o heteróloga respecto al hospedante, o podría ser una secuencia natural o una secuencia sintética. Por "extraña", se entiende que la región de iniciación transcripcional no se da en el hospedante natural en el cual se introduce la región de iniciación transcripcional. Por "promotor funcional" se entiende el promotor, que cuando está ligado operativamente a una secuencia que codifica una proteína de interés, es capaz de conducir la expresión (es decir, la transcripción y la traducción) de la proteína codificada, o, cuando está ligado operativamente a una secuencia inhibidora que codifica una molécula de nucleótido inhibidora (por ejemplo, un ARN de horquilla, un ARN de cadena doble, un polinucleótido de miARN, y similares), el promotor es capaz de iniciar la transcripción de la secuencia inhibidora ligada operativamente de tal modo que se expresa la molécula de nucleótido inhibidora. Los promotores se pueden seleccionar en base al resultado deseado. Así, las casetes de expresión pueden comprender promotores constitutivos, de tejido preferido u otros promotores para la expresión en plantas.

Tal como se usa en la presente memoria, un gen quimérico comprende una secuencia codificadora ligada operativamente a una región de iniciación de la transcripción que es heteróloga respecto a la secuencia codificadora.

Se puede emplear cualquier promotor adecuado conocido en la técnica según la presente invención, incluyendo promotores bacterianos, levaduras, fúngicos, de insecto, de mamífero y vegetales. Por ejemplo, se puede usar promotores vegetales, que incluyen promotores de lenteja de agua. Los ejemplos de promotores incluyen, aunque sin limitación, el promotor 35S de Virus de Mosaico de Coliflor, los promotores de opina sintetasa (por ejemplo, nos, mas, ocs, etc.), el promotor de ubiquitina, el promotor de actina, el promotor de subunidad pequeña de ribulosa bifosfato (RubP) carboxilasa, y el promotor de alcohol deshidrogenasa. El promotor de subunidad pequeña de RubP

carboxilasa de lenteja de agua es conocido en la técnica (Silverthorne y col. (1990) Plant Mol. Biol. 15: 49). Otros promotores procedentes de virus que infectan plantas, preferiblemente lenteja de agua, también son adecuados e incluyen, aunque sin limitación, promotores de virus de mosaico de Dasheen, de virus de Chlorella (por ejemplo, promotor de adenina metiltransferasa de virus de Chlorella; Mitra y col. (1994) Plant Mol. Biol. 26: 85), virus de marchitez moteada del tomate, virus cascabel del tabaco, virus de necrosis del tabaco, virus de mancha anular del tabaco, virus de mancha anular del tomate, virus de mosaico de pepino, virus del enanismo de cacahuete, virus de mosaico de la alfalfa, badnavirus baciliforme de la caña de azúcar y otros similares.

Otros elementos de control de la expresión adecuados se describen en la solicitud provisional de patente en tramitación con la presente y poseída en común con la presente titulada "Expression Control Elements from the Lemnaceae Family", asignada a la Solicitud de Patente de EE.UU. N° 60/759.308, presentada el 17 de enero de 2006 (ver también el documento WO2007124186). Los elementos de control de la expresión descritos en dicha solicitud en tramitación con la presente se aislaron a partir de genes de ubiquitina para varios miembros de la familia de las Lemnaceae, y por tanto se denominan "elementos de control de expresión de ubiquitina de Lemnaceae". La SEC ID N°: 7 descrita en la presente memoria establece el elemento de control de la expresión de ubiquitina de *Lemna minor* de longitud completa, que incluye el promotor más 5' UTR (nucleótidos 1-1625) y el intrón (nucleótidos 1626-2160). La SEC ID N°: 8 establece el elemento de control de la expresión de ubiquitina de *Spirodella polyrrhiza* de longitud completa, que incluye el promotor más 5' UTR (nucleótidos 1-1041) y el intrón (nucleótidos 1042-2021). La SEC ID N°: 9 establece el elemento de control de la expresión de ubiquitina de *Lemna aequinoctialis* de longitud completa, que incluye el promotor más 5' UTR (nucleótidos 1-964) y el intrón (nucleótidos 965-2068). La SEC ID N°: 10 establece el promotor más la porción 5' UTR del elemento de control de la expresión de ubiquitina de *L. minor* (designado en la presente memoria "promotor de LmUbq"). La SEC ID N°: 11 establece el promotor más la porción 5' UTR del elemento de control de la expresión de ubiquitina de *S. polyrrhiza* (designado en la presente memoria "promotor SpUbq"). La SEC ID N°: 12 establece el promotor más la porción 5' UTR del elemento de control de la expresión de ubiquitina de *L. aequinoctialis* (designado en la presente memoria "promotor LaUbq"). La SEC ID N°: 13 establece la porción de intrón del elemento de control de la expresión de ubiquitina de *L. minor* (designado en la presente memoria "promotor LmUbq"). La SEC ID N°: 14 establece la porción de intrón del elemento de control de la expresión de ubiquitina de *S. polyrrhiza* (designado en la presente memoria "intrón LaUbq"). Se reconoce que el promotor individual más las secuencias 5' UTR establecidas en las SEC ID N°: 10-12, y las variantes y fragmentos biológicamente activos de los mismos, se pueden usar para regular la transcripción de las secuencias de nucleótidos ligadas operativamente de interés en plantas. De forma similar, una o más de las secuencias de intrón establecidas en las SEC ID N°: 13-15, y los fragmentos o variantes biológicamente activos de las mismas, se pueden ligar operativamente a un promotor de interés, que incluye un promotor establecido en la SEC ID N°: 10, 11 ó 12 con el objetivo de potenciar la expresión de una secuencia de nucleótidos que esté ligada operativamente a dicho promotor.

Los fragmentos y las variantes de los elementos de control de la expresión también se pueden usar dentro de las casetes de expresión para dirigir la expresión del polinucleótido ligado operativamente de interés. Por "fragmento de un elemento de control de la expresión" se entiende una porción del elemento de control de la expresión de longitud completa, tal como una porción de uno cualquiera de los elementos de control de la expresión establecidos en la SEC ID N°: 7-9. Los fragmentos de un elemento de control de la expresión retienen la actividad biológica y por tanto abarcan fragmentos capaces de iniciar o potenciar la expresión de un polinucleótido ligado operativamente de interés. Por tanto, por ejemplo, se puede utilizar menos que los elementos de control de la expresión enteros descritos en la presente memoria para conducir la expresión de un polinucleótido ligado operativamente de interés. Los ejemplos específicos, no limitantes, de dichos fragmentos de un elemento de control de la expresión incluyen las secuencias de nucleótidos establecidas en una cualquiera de las SEC ID N°: 10-12 (tal como se ha descrito anteriormente en la presente memoria), así como las truncamientos 5' del elemento de control de la expresión de ubiquitina de *L. minor* (SEC ID N°: 7), tal como los nucleótidos 1288-2160 de SEC ID N°: 7 (promotor truncado LmUbq N° 1) y los nucleótidos 1132-2160 de SEC ID N°: 1 (promotor truncado LmUbq N°: 2).

Los nucleótidos de dichos fragmentos normalmente comprenderán la secuencia de reconocimiento TATA del elemento de control de la expresión particular. Dichos fragmentos pueden obtenerse mediante el uso de enzimas de restricción para romper los elementos de control de la expresión naturales descritos en la presente memoria; sintetizando una secuencia de nucleótidos a partir de la secuencia natural de la secuencia de ADN del elemento de control de la expresión; o pueden obtenerse a través del uso de tecnología de reacción en cadena de polimerasa (PCR). Véase particularmente, Mullis et al. (1987) Methods Enzymol. 155: 335-350, y Erlich, ed. (1989) PCR Technology (Stockton Press, Nueva York).

Las variantes de los elementos de control de la expresión, tales como aquellas que son el resultado de mutagénesis dirigida a sitio, también se pueden usar en las casetes de expresión para proporcionar la expresión del polinucleótido ligado operativamente de interés. Por "variante de un elemento de control de la expresión" se entiende secuencias que tengan una similitud sustancial con un elemento de control de la expresión descrito en la presente memoria (por ejemplo, el elemento de control de la expresión establecido en las SEC ID N°: 7, 8 ó 9), o con un fragmento del mismo (por ejemplo, las respectivas secuencias establecidas en las SEC ID N°: 10-15). Las variantes naturales de elementos de control de la expresión se pueden identificar con el uso de técnicas de biología molecular bien conocidas, tal como, por ejemplo, con técnicas de PCR o de hibridación como las descritas anteriormente. Las variantes de elementos de control de la expresión también incluyen secuencias de nucleótidos derivados

5 sintéticamente tales como los generados, por ejemplo, usando mutagénesis dirigida a sitio. Generalmente, las variantes de un elemento de control de la expresión particular descrito en la presente memoria, que incluyen las variantes de las SEC ID N°: 7-15, tendrán una identidad de secuencia de al menos un 40%, 50%, 60%, 65%, 70%, generalmente al menos un 75%, 80%, 85%, preferiblemente aproximadamente un 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97% y más preferiblemente de aproximadamente 98%, 99% o más, con respecto a dicha secuencia de nucleótidos particular, determinada mediante programas de alineamiento de secuencia descritos anteriormente en la presente memoria usando los parámetros por defecto.

10 Los elementos de control de la expresión, que incluyen promotores, pueden seleccionarse para producir un nivel de regulación deseado. Por ejemplo, en algunos casos, puede ser ventajoso usar un promotor que confiera una expresión constitutiva (por ejemplo, el promotor de manopina sintasa de *Agrobacterium tumefaciens*). Alternativamente, en otras situaciones, por ejemplo, en lo que respecta a la expresión de un polipéptido heterólogo, puede ser ventajoso usar promotores que sean activados como respuesta a estímulos ambientales específicos (por ejemplo, promotores génicos de choque térmico, promotores génicos inducibles por arrastre, promotores génicos inducibles por patógeno, promotores génicos inducibles por lesión y promotores génicos inducibles por luz/oscuridad) o reguladores de crecimiento de plantas (por ejemplo, promotores de genes inducidos por ácido abscísico, auxinas, citokininas y ácido giberélico). Como alternativa adicional, los promotores se pueden seleccionar para producir expresión específica de tejido (por ejemplo, promotores específicos de raíz, hoja y flores).

20 La fuerza global de un promotor dado puede verse influido por la combinación y la organización espacial de las secuencias de nucleótidos que actúan en cis tal como las secuencias activadoras por encima (en dirección 5'). Por ejemplo, las secuencias de nucleótidos activadoras derivadas del gen de octopina sintasa de *Agrobacterium tumefaciens* pueden potenciar la transcripción del promotor de manopina sintasa de *Agrobacterium tumefaciens* (véase la Patente de EE.UU. N° 5.955.646 a nombre de Gelvin et al.). En la presente invención, la casete de expresión puede contener secuencias de nucleótido activadoras insertadas por encima (en dirección 5') de la secuencia de promotor para potenciar la expresión de la secuencia de nucleótidos de interés. En una realización, la casete de expresión incluye tres secuencias activadoras por encima (en dirección 5') derivadas del gen de octopina sintasa de *Agrobacterium tumefaciens* ligadas operativamente a un promotor derivado de un gen de manopina sintasa de *Agrobacterium tumefaciens* (ver la Patente de EE.UU. N° 5.955.646).

30 Cuando el elemento de control de la expresión se vaya a usar para conducir la expresión de una secuencia de ADN ligada operativamente que codifica una molécula de ARNhp pequeña, por ejemplo, dentro de una casete de expresión de ARNi descrita anteriormente en la presente memoria, es ventajoso usar un elemento de control de la expresión que comprende un promotor reconocido por la ARN polimerasa III dependiente de ADN. Tal como se usa en la presente memoria, "un promotor reconocido por la ARN polimerasa III dependiente de ADN" es un promotor que dirige la transcripción de la región de ADN asociada a través de la acción de polimerasa de la ARN polimerasa III. Esto incluye los genes que codifican 5S ARN, tARN, 7SL ARN, U6 snARN y otros ARNs estables pequeños, muchos implicados en el procesamiento de ARN. La mayoría de los promotores usados por Pol III requieren los elementos de secuencia por debajo (dirección 3') de +1, dentro de la región transcrita. Sin embargo, una minoría de las plantillas de pol III carece de cualquier requerimiento para elementos de promotor intragénicos. Estos se denominan promotores de tipo 3. Por "promotores de Pol III de tipo 3" se entiende aquellos promotores que son reconocidos por ARN polimerasa III y que contienen elementos que actúan en cis, que interactúan con la ARN polimerasa III por encima (en dirección 5') de la región normalmente transcrita por la ARN polimerasa III. Dichos promotores de Pol III de tipo 3 pueden ensamblarse dentro de las casetes de expresión de ARNi de la invención para conducir la expresión de la secuencia de ADN ligada operativamente que codifica la molécula de ARNhp pequeña.

45 Normalmente, los promotores de Pol III de tipo 3 contienen una caja TATA (localizada entre -25 y -30 en el gen U6 snARN humano) y un elemento de Secuencia Proximal (PSE; localizado entre -47 y -66 en U6 snARN humano). También pueden contener un Elemento de Secuencia Distal (DSE; localizado entre -214 y -244 en el U6 snARN humano). Los promotores de Pol III de tipo 3 se pueden encontrar, por ejemplo, asociados a los genes que codifican 7SL ARN, U3 snARN y U6 snRNA. Dichas secuencias han sido aisladas a partir de *Arabidopsis*, arroz y tomate. Véase, por ejemplo, las SEC ID N°: 1-8 de la Publicación de Solicitud de Patente de EE.UU. N°: 20040231016.

50 Otras secuencias de nucleótidos para los promotores de Pol III de tipo 3 pueden encontrarse en las bases de datos de secuencias de nucleótidos bajo las entradas para los genes de *A. thaliana* AT7SL-1 para 7SL ARN (X72228), gen de *A. thaliana* AT7SL-2 para 7SL ARN (X72229), gen de *A. thaliana* AT7SL-3 para 7SL ARN (AJ290403), gen de *Humulus lupulus* H17SL-1 (AJ236706), gen de *Humulus lupulus* H17SL-2 (AJ236704), gen de *Humulus lupulus* H17SL-3 (AJ236705), gen de *Humulus lupulus* H17SL-4 (AJ236703), gen de *A. thaliana* U6-1 snRNA (X52527), gen de *A. thaliana* U6-26 snRNA (X52528), gen de *A. thaliana* U6-29 snRNA (X52529), gen de *A. thaliana* U6-1 snRNA (X52527), gen de *Zea mays* U3 snRNA (Z29641), gen de *Solanum tuberosum* U6 snARN (Z17301; X 60506; S83742), gen de tomate U6 ARN nuclear pequeño (X51447), gen de *A. thaliana* U3C snARN (X52630), gen de *A. thaliana* U3B snARN (X52629), promotor de *Oryza sativa* U3 snARN (X79685), gen de tomate U3 ARN nuclear pequeño (X14411), gen de *Triticum aestivum* U3 snARN (X63065) y gen de *Triticum aestivum* U6 snARN (X63066).

60 Se pueden aislar otros promotores de Pol III de tipo 3 a partir de variedades de tomate, arroz o *Arabidopsis*, o a partir de otras especies vegetales usando métodos bien conocidos en la técnica. Por ejemplo, se pueden aislar

bibliotecas de clones genómicos a partir de dichas plantas usando como sonda secuencias codificadoras de U6 snARN, U3 snARN ó 7SL ARN (tales como las secuencias codificadoras de cualquiera de las secuencias mencionadas anteriormente identificadas por su número de acceso, y adicionalmente la secuencia codificadora de U6 snARN de *Vicia faba* (X04788), el ADN de maíz para U6 snARN (X52315) o el ADN de maíz para 7SL ARN (X14661)), y las secuencias por encima (en dirección 5'), preferiblemente los aproximadamente de 300 a 400 pb por encima de las regiones transcritas, se pueden aislar y usarse como promotores Pol III de tipo 3. Alternativamente, se pueden usar técnicas basadas en PCR tales como la PCR inversa o la TAILTM-PCR para aislar las secuencias genómicas que incluyen las secuencias promotores adyacentes a regiones transcritas conocidas. Además, cualquiera de las secuencias de promotor de Pol III de tipo 3 descritos en la presente memoria, identificados por sus números de acceso y SEC ID N°, pueden usarse como sondas en condiciones de hibridación severas o como fuente de información para generar cebadores de PCR para aislar las correspondientes secuencias promotoras a partir de otras variedades o especies de plantas.

Aunque los promotores de Pol III de tipo 3 no tienen ningún requisito para elementos que actúan en cis localizados con la región transcrita, no obstante es evidente que las secuencias localizadas normalmente por debajo del sitio de iniciación de la transcripción pueden incluirse en las casetes de expresión de ARNi. Además, aunque los promotores de Pol III de tipo 3 aislados originalmente a partir de plantas monocotiledóneas pueden usarse efectivamente en casetes de expresión de ARNi para suprimir la expresión de un gen diana tanto en células y plantas dicotiledóneas como monocotiledóneas, los promotores de Pol III de tipo 3 aislados originalmente a partir de plantas dicotiledóneas solo pueden usarse eficientemente (según se ha publicado) en células y plantas dicotiledóneas. Además, el gen más efectivo para silenciar (según se ha publicado) se obtiene cuando la casete de expresión se diseña para comprender un promotor de Pol III de tipo 3 derivado de la misma especie o de una estrechamente relacionada. Véase, por ejemplo, la Publicación de Solicitud de Patente de EE.UU. n° 20040231016. Así, cuando la planta de interés es una planta monocotiledónea, y el método elegido para inhibir la expresión de FucT y/o XylT es la interferencia de ARNhp pequeño, el promotor de Pol III de tipo 3 preferiblemente procede de otra planta monocotiledónea, que incluye la especie vegetal para la cual se va a alterar el modelo de glicosilación de los glicanos *N*-ligados de una glicoproteína de interés.

La casete de expresión incluye por tanto en la dirección 5'-3' de la transcripción, un elemento de control de la expresión que comprende una región de iniciación transcripcional y traduccional, un polinucleótido de interés, por ejemplo, una secuencia que codifica una proteína heteróloga de interés o una secuencia que codifica una secuencia inhibidora de FucT o XylT que, cuando se expresa, es capaz de inhibir la expresión o la función de FucT y/o XylT, y una región de terminación transcripcional y traduccional funcional en plantas. Se puede usar cualquier secuencia de terminación adecuada conocida en la técnica según la presente invención. La región de terminación puede ser nativa con la región de iniciación transcripcional, puede ser nativa con la secuencia de nucleótidos de interés o puede derivarse de cualquier otra fuente. Las regiones de terminación convenientes están disponibles a partir del plásmido-Ti de *A. tumefaciens*, tal como las regiones de octopina sintetasa y la nopalina sintetasa. Véase también Guerineau et al. (1991) Mol. Gen. Genet. 262:141; Proudfoot (1991) Cell 64:671; Sanfacon et al. (1991) Genes Dev. 5: 141; Mogen et al. (1990) Plant Cell 2: 1261; Munroe et al. (1990) Gene 91: 151; Ballas et al. (1989) Nucleic Acids Res. 17: 7891 y Joshi et al. (1987) Nucleic Acids Res. 15: 9627. Los ejemplos adicionales de secuencias de terminación son la secuencia de terminación de subunidad pequeña de RubP carboxilasa de guisante y la secuencia de terminación de virus 35S de Mosaico de Coliflor. Otras secuencias de terminación adecuadas serán evidentes para los especialistas en la técnica, e incluyen la tira oligo dT descrita en la presente memoria anteriormente para uso con promotores de Pol III de tipo 3 que conduce la expresión de un polinucleótido inhibidor de FucT y/o XylT que forma una estructura de ARNhp.

Alternativamente, el(los) polinucleótido(s) de interés pueden proporcionarse en cualquier casete de expresión adecuada conocida en la técnica.

Generalmente, la casete de expresión comprenderá un gen marcador seleccionable para la selección de células o tejidos transformados. Los genes marcadores seleccionables incluyen genes que codifican resistencia a antibióticos, tal como los que codifican neomicina fosfotransferasa II (NEO) e higromicina fosfotransferasa (HPT), así como los genes que confieren resistencia a compuestos herbicidas. Los genes de resistencia a herbicidas generalmente codifican para una proteína diana modificada insensible al herbicida o para una enzima que degrada o detoxifica el herbicida en la planta antes de que pueda actuar. Véase DeBlock et al. (1987) EMBO J. 6: 2513; DeBlock et al. (1989) Plant Physiol. 91: 691; Fromm et al. (1990) BioTechnology 8: 833; Gordon-Kamm et al. (1990) Plant Cell 2: 603. Por ejemplo, se ha obtenido la Resistencia a herbicidas de glifosfato o sulfonilurea usando genes que codifican para las enzimas diana mutantes, 5-enolpiruvilsikimato-3-fosfato sintasa (EPSPS) y acetolactato sintasa (ALS). Se ha obtenido resistencia a amonio de glufosinato, boromoxinilo y 2,4-diclorofenoxiacetato (2,4-D) usando genes bacterianos que codifican fosfinotricina acetiltransferasa, una nitrilasa, o una mono-oxigenasa de 2,4-diclorofenoxiacetato mono-oxigenasa, que detoxifica los respectivos herbicidas.

Para los propósitos de la presente invención, los genes marcadores seleccionables incluyen, aunque sin limitación, genes que codifican la neomicina fosfotransferasa II (Fraley y col. (1986) CRC Critical Reviews in Plant Science 4: 1); cianimida hidratasa (Maier-Greiner et al. (1991) Proc. Natl. Acad. Sci. USA 88: 4250); aspartato quinasa; dihidrodipicolinato sintasa (Perl et al. (1993) BioTechnology 11: 715); gen bar (Toki et al. (1992) Plant Physiol. 100: 1503; Meagher et al. (1996) Crop Sci. 36:1367); triptófano descarboxilasa (Goddijn et al. (1993) Plant Mol. Biol. 22:

- 907); neomicina fosfotransferasa (NEO; Southern et al. (1982) J. Mol. Appl. Gen. 1: 327); higromicina fosfotransferasa (HPT ó HYG; Shimizu et al. (1986) Mol. Cell. Biol. 6: 1074); dihidrofolato reductasa (DHFR; Kwok et al. (1986) Proc. Natl. Acad. Sci. USA 83: 4552); fosfinotricina acetiltransferasa (DeBlock et al. (1987) EMBO J. 6: 2513); ácido 2,2-dicloropropiónico deshalogenasa (Buchanan-Wollaston et al. (1989) J. Cell. Biochem. 13D: 330); acetohidroxiácido sintasa (Patente de EE.UU. Nº 4.761.373 a nombre de Anderson et al.; Haughn et al. (1988) Mol. Gen. Genet. 221: 266); 5-enolpiruvil-sikimato-fosfato sintasa (aroA; Comai et al. (1985) Nature 317: 741); haloarilnitrilasa (documento WO 87/04181 a nombre de Stalker et al.); acetil-coenzima A carboxilasa (Parker et al. (1990) Plant Physiol. 92: 1220); dihidropteroata sintasa (sull; Guerinneau et al. (1990) Plant Mol. Biol. 15: 127); y polipéptido de fotosistema II de 32 kDa (psbA; Hirschberg et al. (1983) Science 222: 1346 (1983).
- 10 También se incluyen los genes que codifican resistencia a: gentamicina (por ejemplo, aacC1, Wohlleben et al. (1989) Mol. Gen. Genet. 217: 202-208); cloranfenicol (Herrera-Estrella et al. (1983) EMBO J. 2: 987); metotrexato (Herrera-Estrella et al. (1983) Nature 303: 209; Meijer et al. (1991) Plant Mol. Biol. 16: 807); higromicina (Waldron et al. (1985) Plant Mol. Biol. 5: 103; Zhijian et al. (1995) Plant Science 108: 219; Meijer et al. (1991) Plant Mol. Biol. 16: 807); estreptomycin (Jones et al. (1987) Mol. Gen. Genet. 210: 86); espectinomycin (Bretagne-Sagnard et al. (1996) Transgenic Res. 5: 131); bleomicina (Hille et al. (1986) Plant Mol. Biol. 7: 171); sulfonamida (Guerinneau et al. (1990) Plant Mol. Biol. 15: 127); bromoxinil (Stalker et al. (1988) Science 242: 419); 2,4-D (Streber et al. (1989) BioTechnology 7: 811); fosfinotricina (DeBlock et al. (1987) EMBO J. 6: 2513); espectinomycin (Bretagne-Sagnard y Chupeau, Transgenic Research 5: 131).
- 15 El gen bar confiere resistencia a herbicida frente a los herbicidas de tipo glufosinato, tal como la fosfinotricina (PPT) o el bialafos, y otros similares. Como se ha indicado anteriormente, otros marcadores seleccionables que podrían usarse en las construcciones de vector incluyen, aunque sin limitación el gen pat, también para resistencia a bialafos y fosfinotricina, el gen ALS para resistencia a imidazolina, el gen HPH o HYG para resistencia a higromicina, el gen de EPSP sintasa para resistencia a glifosato, el gen Hm1 para resistencia a la toxina-Hc, y otros agentes selectivos usados rutinariamente y conocidos por el especialista en la técnica. Véase Yarranton (1992) Curr. Opin. Biotech. 3: 506; Christopherson et al. (1992) Proc. Natl. Acad. Sci. USA 89: 6314; Yao et al. (1992) Cell 71: 63; Reznikoff (1992) Mol. Microbiol. 6: 2419; Barkley et al. (1980) The Operon 177-220; Hu et al. (1987) Cell 48: 555; Brown et al. (1987) Cell 49: 603; Figge et al. (1988) Cell 52: 713; Deuschle et al. (1989) Proc. Natl. Acad. Sci. USA 86: 5400; Fuerst et al. (1989) Proc. Natl. Acad. Sci. USA 86: 2549; Deuschle et al. (1990) Science 248: 480; Labow et al. (1990) Mol. Cell. Biol. 10: 3343; Zambretti et al. (1992) Proc. Natl. Acad. Sci. USA 89: 3952; Baim et al. (1991) Proc. Natl. Acad. Sci. USA 88: 5072; Wyborski et al. (1991) Nuc. Acids Res. 19: 4647; Hillenand-Wissman (1989) Topics in Mol. And Struc. Biol. 10: 143; Degenkolb et al. (1991) Antimicrob. Agents Chemother. 35: 1591; Kleinschmidt et al. (1988) Biochemistry 27: 1094; Gatz et al. (1992) Plant J. 2: 397; Gossen et al. (1992) Proc. Natl. Acad. Sci. USA 89: 5547; Oliva et al. (1992) Antimicrob. Agents Chemother. 36: 913; Hlavka et al. (1985) Handbook of Experimental Pharmacology 78; y Gill et al. (1988) Nature 334: 721.
- 20 La lista anterior de genes marcadores seleccionables no pretende ser limitativa. En la presente invención se puede usar cualquier gen marcador seleccionable.

Modificación de Secuencias de Nucleótidos para Expresión Potenciada en una Planta Hospedante:

- 40 Cuando la planta de interés también está modificada genéticamente para expresar una proteína heteróloga de interés, por ejemplo, una planta hospedante transgénica que actúa como sistema de expresión para la producción recombinante de una proteína heteróloga, la presente invención proporciona para la modificación de la secuencia de polinucleótido expresada que codifica la proteína heteróloga de interés para potenciar su expresión en la planta hospedante. Por tanto, cuando sea apropiado, los polinucleótidos pueden optimizarse para aumentar la expresión en la planta transformada. Es decir, los polinucleótidos pueden sintetizarse usando codones preferidos de planta para una expresión mejorada. Véase, por ejemplo, Campbell y Gowri (1990) Plant Physiol. 92: 1-11 para una discusión de uso de codones de hospedante preferido. En la técnica se dispone de métodos para sintetizar secuencias de nucleótidos con codones preferidos de planta. Véase, por ejemplo, las Patentes de EE.UU. Nº 5.380.831 y 5.436.391; Perlak et al. (1991) Proc. Natl. Acad. Sci. USA 15: 3324; Iannacome et al. (1997) Plant Mol. Biol. 34: 485; y Murray et al., (1989) Nucleic Acids. Res. 17: 477.
- 45 En esta invención, la planta hospedante es un miembro de la familia de las lentejas de agua, y el polinucleótido que codifica el polipéptido heterólogo de interés, por ejemplo, un polipéptido de mamífero, se modifica para potenciar la expresión del polipéptido heterólogo codificado. De esta manera, dicha modificación es la síntesis del polinucleótido que codifica el polipéptido heterólogo de interés usando codones preferidos de lenteja de agua, en donde la síntesis se puede llevar a cabo usando cualquier método conocido por el especialista en la técnica. Los codones preferidos pueden determinarse a partir de los codones de mayor frecuencia en las proteínas expresadas en lenteja de agua.
- 50 Por ejemplo, la frecuencia de uso de codón para la *Lemna gibba* se puede encontrar en la página web: [http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+gibba+\[gbpln\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+gibba+[gbpln]), y la frecuencia de uso de codón para la *Lemna minor* se puede encontrar en la página web [http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+minor+\[gbpln\]](http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=Lemna+minor+[gbpln]) y en la Tabla 1. Se reconoce que los genes heterólogos que han sido optimizados para expresión en lenteja de agua y otras monocotiledóneas, así como otras dicotiledóneas, se pueden usar en los métodos de la invención. Véase, por ejemplo los documentos EP 0 359 472, EP 0 385 962, WO 91/16432; Perlak et al. (1991) Proc. Natl. Acad. Sci. USA 88: 3324; Iannacome et

al. (1997) Plant Mol. Biol. 34: 485; and Murray et al. (1989) Nuc. Acids Res. 17:477. Se reconoce además que parte o todo el nucleótido que codifica el polipéptido heterólogo de interés puede optimizarse o puede ser sintético. En otras palabras, también se pueden usar secuencias parcial o totalmente optimizadas. Por ejemplo, el 40 %, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 87%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% el 100% de los codones pueden ser codones preferidos de lenteja de agua. En una realización, entre el 90 y el 96% de los codones son codones preferidos de lenteja de agua. La secuencia codificadora de una secuencia de polinucleótido que codifica un polipéptido heterólogo de interés puede comprender codones usados con una frecuencia de al menos el 17% en *Lemna gibba*. En una realización, la secuencia de nucleótidos modificada es el α -2B-interferón humano que codifica la secuencia de nucleótido mostrada en la SEC ID N°: 16, que contiene un 93% de codones preferidos de lenteja de agua.

Tabla 1: codones preferidos de Lemna gibba del GenBank Release 113

UUU 2.2	(4)	UCU 0.5	(1)	UAU 2.2	(4)	UGU 0.0	(0)
UUC 50.5	(92)	UCC 31.9	(58)	UAC 40.1	(73)	UGC 17.6	(32)
UUA 0.0	(0)	UCA 0.5	(1)	UAA 3.8	(7)	UGA 1.6	(3)
UUG 2.7	(5)	UCG 15.4	(28)	UAG 0.0	(0)	UGG 24.2	(44)
CUU .0.5	(1)	CCU 6.6	(12)	CAU 0.5	(1)	CGU 1.1	(2)
CUC 39.0	(71)	CCC 43.4	(79)	CAC 6.6	(12)	CGG 36.9	(49)
CUA 1.1	(2)	CHCA 2.2	(4)	CAA 4.4	(8)	CGA 1.1	(2)
CUG 22.5	(41)	CCG 20.9	(38)	CAG 26.9	(49)	CGG 7.7	(14)
AAU 0.0	(0)	ACU 3.3	(6)	AAU 1.1	(2)	AGU 0.0	(0)
AUC 33.5	(61)	ACC 26.4	(48)	AAC 37.9	(69)	AGC 22.0	(40)
AUA 0.0	(0)	ACA 0.5	(1)	AAA 0.0	(0)	AGA 4.9	(9)
AUG 33.5	(61)	ACG 9.3	(17)	AAG 57.1	(104)	AGG 6.0	(11)
GUU 9.3	(17)	GCU 7.1	(13)	GAU 1.6	(3)	GGU 1.1	(2)
GUC 28.0	(51)	GCC 73.6	(134)	GAC 38.4	(70)	GGC 46.7	(85)
GUA 0.0	(0)	GCA 5.5	(10)	GAA 2.2	(4)	GGA 1.1	(2)
GUG 34.0	(62)	GCG 20.9	(38)	GAG 62.6	(114)	GGG 27.5	(50)

También se pueden hacer otras modificaciones al polinucleótido que codifica el polipéptido heterólogo de interés para potenciar su expresión en una planta hospedante de interés, que incluye la lenteja de agua. Estas modificaciones incluyen, aunque sin limitación, la eliminación de secuencias que codifican señales de poliadenilación espurias, señales de sitio de división exón-intrón, repeticiones de tipo transposón, y otras secuencias bien caracterizadas de ese tipo que pueden ser perjudiciales para la expresión génica. El contenido G-C de la secuencia se puede ajustar a los niveles medios para un hospedante celular dado, calculados mediante referencia a genes conocidos expresados en la célula hospedante. Cuando sea posible, el polinucleótido que codifica el polipéptido heterólogo de interés puede modificarse para evitar las estructuras de ARNm secundarias de horquilla predichas.

Existen diferencias conocidas entre las secuencias de nucleótidos de contexto de iniciación de traducción óptima para los codones de iniciación de la traducción de animales y de plantas, y la composición de dichas secuencias de nucleótidos de contexto de iniciación de traducción puede influir en la eficacia de la iniciación de la traducción. Véase, por ejemplo, Lukaszewicz et al. (2000) Plant Science 154: 89-98; y Joshi et al. (1997); Plant Mol. Biol. 35: 993-1001. En la presente invención, la secuencia de nucleótidos de contexto de iniciación de la traducción para el codón de iniciación de la traducción del polinucleótido de interés, por ejemplo, el polinucleótido que codifica un polipéptido heterólogo de interés, puede modificarse para potenciar la expresión en la lenteja de agua. En una realización, la secuencia de nucleótidos se modifica de tal modo que los tres nucleótidos directamente por encima

(dirección 5') del codón de iniciación de la traducción de la secuencia de nucleótidos de interés son "ACC". En una segunda realización, dichos nucleótidos son "ACA".

La expresión de un transgén en una planta hospedante, que incluye una lenteja de agua, también se puede potenciar mediante el uso de secuencias líder 5'. Dichas secuencias líder pueden actuar para potenciar la traducción. En la técnica se conocen líderes de traducción e incluyen, aunque sin limitación, líderes de picoravirus, por ejemplo, líder EMCV (región no codificadora 5' de encefalomiocarditis; Elroy-Stein et al. (1989) Proc. Natl. Acad. Sci USA 86: 6126); líderes de potivirus, por ejemplo, líder TEV (Virus de Grabado de Tabaco; Allison et al. (1986) Virology 154: 9); proteína de unión de cadena pesada de inmunoglobulina humana (BiP; Macajak y Sarnow (1991) Nature 353: 90); líder no traducido de proteína de recubrimiento de ARNm de virus de mosaico de alfalfa (AMV RNA 4; Jobling y Gehrke (1987) Nature 325: 622); líder de virus de mosaico de tabaco (TMV; Gallie (1989) Molecular Biology of RNA, 23: 56); líder de virus de grabado de patata (Tomashevskaya et al. (1993) J. Gen. Virol. 74: 2717-2724); región no traducida 5' Fed-1 (Dickey (1992) EMBO J. 11: 2311-2317); región no traducida 5' RbcS (Silverthorne et al. (1990) J. Plant. Mol. Biol. 15: 49-58); y líder de virus de moteado clorótico de maíz (MCMV; Lommel et al. (1991) Virology 81: 382). Véase también, Della-Cioppa et al. (1987) Plant Physiology 84: 965. La secuencia líder comprende la secuencia de intrón vegetal, que incluye la secuencia de intrón del gen de alcohol deshidrogenasa 1 (ADH1) de maíz, el gen de catalasa de semilla de ricino, o el gen PAT1 de mecanismo triptófano de Arabidopsis, también han demostrado aumentar la eficacia traduccional en plantas (Callis et al. (1987) Genes Dev. 1: 1183-1200; Mascarenhas et al. (1990) Plant Mol. Biol. 15: 913-920). Véase también el documento 2007/0180583, en donde la secuencia líder que comprende una secuencia de intrón de lenteja de agua seleccionada del grupo que consiste en los intrones establecidos en las SEC ID N°: 13-15 proporciona una mayor eficiencia traduccional en la lenteja de agua.

En algunas realizaciones de la presente invención, la secuencia de nucleótidos correspondiente a los nucleótidos 1222-1775 del gen de alcohol deshidrogenasa 1 de maíz (ADH1; Número de Acceso GenBank X04049), o la secuencia de nucleótidos correspondiente al intrón establecido en las SEC ID N°: 13, 14 ó 15, se inserta por encima (en dirección 5') del polinucleótido que codifica el polipéptido heterólogo de interés o el polinucleótido inhibidor de FucT y/o XylT para potenciar la eficacia de su traducción. En otra realización, la casete de expresión contiene el líder del gen de subunidad 5B pequeña de ribulosa-bis-fosfato carboxilasa de *Lemna gibba* (líder RbcS; véase Buzby y col. (1990) Plant Cell 2: 805-814; véase también las SEC ID N°: 16, 17 ó 18 de la presente invención).

Véase también, solo a modo de ejemplo, los vectores de expresión descritos en las figuras de la presente memoria, en donde el líder RbcS y el intrón ADH1 se incluyen en las secuencias reguladoras por encima (en dirección 5') dentro de una casete de expresión que comprende el polinucleótido inhibidor de FucT (Figura 8), el polinucleótido inhibidor de XylT (Figuras 9 y 11), una casete de expresión que comprende la molécula inhibidora quimérica de FucT/XylT (Figura 10), o una casete de expresión que comprende la secuencia codificadora correspondiente al polipéptido heterólogo, la cadena pesada de IgG 1 de un anticuerpo monoclonal (Figuras 12, 13 y 14) o la cadena ligera de un anticuerpo monoclonal (Figura 14); en donde el promotor LmUbq y el intrón LmUbq se incluyen como secuencias reguladoras hacia arriba (en dirección 5') dentro de la casete de expresión que comprende el polinucleótido inhibidor de FucT (Figura 11), o una casete de expresión que comprende la secuencia codificadora correspondiente al polipéptido heterólogo, la cadena ligera de IgF1 de un anticuerpo monoclonal (Figura 13); en donde el promotor SpUbq y el intrón SpUbq se incluyen como secuencias reguladoras por encima (en dirección 5') dentro de una casete de expresión que comprende el polinucleótido inhibidor de FucT (Figura 13), o una casete de expresión que comprende el polinucleótido inhibidor quimérico de FucT/XylT (Figura 12); y en donde el promotor LaUbq y el intrón LaUbq se incluyen como secuencias reguladoras por encima (en dirección 5') en una casete de expresión que comprende el polinucleótido inhibidor XylT (Figura 13).

Se reconoce que en la presente invención se puede usar cualquier modificación de secuencia de nucleótido potenciadora de la expresión descrita anteriormente, incluyendo cualquier modificación sencilla o cualquier posible combinación de modificaciones. La frase "modificado para una expresión potenciada" en una planta, por ejemplo, una planta de lenteja de agua, tal como se usa en la presente memoria se refiere a una secuencia de polinucleótidos que contiene una cualquiera de las combinaciones de dichas modificaciones.

Péptidos Señal:

Se reconoce que el polipéptido heterólogo de interés puede ser uno que normalmente o de forma ventajosa se expresa como una proteína secretada. Las proteínas secretadas normalmente se traducen a partir de polipéptidos precursores que incluyen un "péptido señal" que interacciona con una proteína receptora sobre la membrana del retículo endoplasmático (RE) para dirigir la traslocalización de la cadena creciente del polipéptido a través de la membrana y hacia el interior del retículo endoplasmático para secreción desde la célula. Dicho péptido señal a menudo se libera del polipéptido precursor para producir un polipéptido "maduro" que carece de péptido señal. En una realización de la presente invención, se expresa un polipéptido biológicamente activo en la planta hospedante de interés, por ejemplo, una lenteja de agua u otra planta superior, a partir de una secuencia de polinucleótido que está ligada operativamente con una secuencia de nucleótidos que codifica un péptido señal que dirige la secreción del polipéptido al medio de cultivo. Los péptidos señal vegetales dirigidos contra la traslocalización de proteínas al retículo endoplasmático (para secreción al exterior de la célula) son conocidos en la técnica. Véase, por ejemplo, la

Patente de EE.UU. N° 6.020.169 a nombre de Lee et al. En la presente invención, se puede usar cualquier péptido señal vegetal para dirigir el polipéptido expresado al RE.

En algunas realizaciones, el péptido señal es el péptido señal de endoquitinasa básica de *Arabidopsis thaliana* (aminoácidos 14-34 del N° de Acceso de Proteína NCBI BAA82823), el péptido señal de extensina (Stiefel et al. (1990) Plant Cell 2: 785-793), el péptido señal de α -amilasa de arroz (aminoácidos 1-31 del N° de Acceso de Proteína NCBI AAA33885), o una secuencia señal de α -amilasa de arroz modificada (SEC ID N°: 17). En otra realización, el péptido señal se corresponde a un péptido señal de una proteína de lenteja de agua secretada.

Alternativamente, se puede usar un péptido señal de mamífero para dirigir polipéptidos recombinantes expresados en una planta modificada genéticamente de la invención, por ejemplo, una lenteja de agua u otra planta superior de interés, para secreción. Se ha demostrado que las células vegetales reconocen péptidos señal de mamífero dirigidos contra el retículo endoplasmático, y que dichos péptidos señal pueden dirigir la secreción de polipéptidos no solo a través de la membrana plasmática sino también a través de la pared celular vegetal. Véanse las Patentes de EE.UU. N° 5.202.422 y 5.639.947 a nombre de Hiatt et al. En una realización de la presente invención, el péptido señal de mamífero dirigido a la secreción de polipéptido es el péptido señal de α -2b-interferón humano (aminoácidos 1-23 del N° de Acceso de Proteína NCBI AAB59402).

En una realización, la secuencia de nucleótidos que codifica el péptido señal se modifica para potenciar la expresión en la planta hospedante de interés, por ejemplo, una lenteja de agua u otra planta superior, utilizando cualquier modificación o combinación de modificaciones descritas anteriormente para la secuencia de polinucleótidos de interés.

El polipéptido biológicamente activo secretado puede ser recolectado a partir del medio de cultivo empleando cualquier medio convencional conocido en la técnica y puede purificarse mediante cromatografía, electroforesis, diálisis, extracción disolvente-disolvente y otros similares. De esta manera, los polipéptidos purificados, tal como se ha definido anteriormente, pueden obtenerse a partir del medio de cultivo.

El sistema hospedante de expresión de proteína es una planta, y el polipéptido secretado biológicamente activo es una glicoproteína de la invención, en donde la glicoproteína tiene un perfil de glicosilación sustancialmente homogéneo, y es sustancialmente homogénea para la glicoforma G0. Cualquier glicoproteína de este tipo que pueda permanecer dentro del material de la planta opcionalmente puede aislarse y purificarse como se ha indicado anteriormente. La glicoproteína secretada puede obtenerse a partir de medio de cultivo de la planta y se puede purificar usando cualquier medio convencional de la técnica tal como se ha indicado anteriormente. De esta manera, la glicoproteína purificada obtenida a partir del material de la planta está sustancialmente libre de material celular de la planta, e incluye realizaciones en las que las preparaciones de glicoproteína tienen menos de aproximadamente 30%, 20%, 10%, 5% ó 1% (en peso seco) de proteínas vegetales contaminantes. En donde la glicoproteína purificada se obtiene a partir del medio de cultivo de la planta, el medio de cultivo de la planta representa menos de aproximadamente el 30%, 20%, 10%, 5% ó 1% (en peso seco) de precursores químicos o productos químicos sin interés proteínico dentro de la preparación de glicoproteínas purificada.

En algunas realizaciones, dichas glicoproteínas purificadas obtenidas a partir de la planta hospedante pueden incluir al menos 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta aproximadamente 30% (en peso seco) de proteína vegetal contaminante. En otras realizaciones, en donde la glicoproteína se recolecta a partir del medio de cultivo de la planta, el medio de cultivo de la planta en dichas glicoproteínas purificadas puede incluir al menos 0,001%, 0,005%, 0,1%, 0,5%, 1%, 1,5%, 2%, 2,5%, 3%, 3,5%, 4%, 4,5%, 5%, 5,5%, 6%, 6,5%, 7%, 7,5%, 8%, 8,5%, 9%, 9,5%, 10%, 15%, 20%, 25% o hasta aproximadamente 30% (en peso seco) de precursores químicos o productos químicos sin interés proteínico dentro de la preparación de glicoproteína purificada. En algunas realizaciones, el aislamiento y la purificación a partir de la planta hospedante, y, cuando es secretada, a partir del medio de cultivo, da como resultado la recuperación de glicoproteína purificada que está libre de proteína vegetal contaminante, libre de componentes del medio de cultivo de plantas y/o libre de ambos, proteínas vegetales contaminantes y componentes del medio de cultivo de plantas.

Cultivos de Plantas Transformadas y Cultivos de Nódulos de Plantas de Lenteja de Agua Transformadas:

Los protocolos de transformación, así como los protocolos para introducir secuencias de nucleótidos en plantas pueden variar dependiendo del tipo de planta o de célula o nódulo vegetal, es decir, monocot o dicto, destinada a transformación. Los métodos adecuados para introducir secuencias de nucleótidos en plantas o células o nódulos vegetales incluyen la microinyección (Crossway et al. (1986) Biotechniques 4: 320-334), electroporación (Riggs et al. (1986) Proc. Natl. Acad. Sci. USA 83: 5602-5606), transformación mediada por *Agrobacterium* (Patentes de EE.UU. N° 5.563.055 y 5.981.840), transferencia génica directa (Paszowski et al. (1984) EMBO J. 3: 2717-2722), aceleración de partículas balísticas (véase, por ejemplo, las Patentes de EE.UU. N° 4.945.050; 5.879.918; 5.886.244; y 5.932.782 y Tomes et al. (1995) "Direct DNA Transfer into Intact Plant Cells via Microprojectile Bombardment," en Plant Cell, Tissue, and Organ Culture: Fundamental Methods, ed. Gamborg y Phillips (Springer-Verlag, Berlin); Mc-Cabe et al. (1988) Biotechnology 6: 923-926). Las células que han sido transformadas pueden

ser cultivadas en plantas según métodos convencionales. Véase, por ejemplo, McCormick et al. (1986) Plant Cell Reports 5: 81-84.

La lenteja de agua transformada establemente utilizada en esta invención puede obtenerse mediante cualquier método conocido en la técnica. En una realización, la lenteja de agua transformada establemente se obtiene mediante uno de los métodos de transferencia génica descritos en la Patente de EE.UU. N° 6.040.498 a nombre de Stomp et al. Dichos métodos incluyen la transferencia génica mediante bombardeo balístico con microproyectiles recubiertos con un ácido nucleico que comprende la secuencia de nucleótidos de interés, transferencia génica mediante electroporación y transferencia génica mediada por *Agrobacterium* que comprende un vector que comprende la secuencia de nucleótidos de interés. En una realización, la lenteja de agua transformada establemente se obtiene a través de uno cualquiera de los métodos mediados por *Agrobacterium* descritos en la Patente de EE.UU. N° 6.040.498 a nombre de Stomp et al. La *Agrobacterium* usada es *Agrobacterium tumefaciens* o *Agrobacterium rhizogenes*.

Se prefiere que las plantas de lenteja de agua transformadas establemente utilizadas en estos métodos exhiban una morfología normal y sean fértiles mediante reproducción sexual. Preferiblemente, las plantas transformadas de la presente invención contienen una copia sencilla del ácido nucleico transferido, y el ácido nucleico transferido no tiene reestructuraciones notables en él. También se prefieren plantas de lenteja de agua en las que el ácido nucleico transferido está presente en bajos números de copias (es decir, no más de cinco copias, alternativamente, no más de tres copias, y como alternativa adicional, menos de tres copias del ácido nucleico por célula transformada).

Los siguientes ejemplos se muestran a modo de ilustración y no como limitación.

Procedimiento experimental

La *Lemna*, una planta acuática pequeña, es una plataforma de expresión escalable y económicamente atractiva para la fabricación de proteínas terapéuticas libres de patógenos humanos y con un claro camino hacia una aprobación reguladora. El sistema de expresión de *Lemna* (LEX SystemSM) permite una expansión clonal rápida de plantas transgénicas, la secreción de proteínas transgénicas, elevados rendimientos de proteína, facilidad de contención que es comparable a sistemas de cultivo de células de mamífero tales como células de CHO, y tiene la ventaja adicional de bajos costes de operación y de capital (Gasdaska et al. (2003) Bioprocessing J. 50-56). Adicionalmente, este sistema de expresión vegetal ofrece la ventaja de un elevado rendimiento de proteínas (en el intervalo de 6-8% de la proteína soluble total (TSP)). Dichos niveles de expresión, en combinación con el elevado contenido proteico de la *Lemna* y su elevada velocidad de crecimiento (tiempo de duplicación de 36 h), permite una producción de > 1 g de mAb por kg de biomasa en un formato robusto y bien controlado.

Ejemplo 1: Aislamiento de Proteínas de *Lemna minor* Implicadas en la N-Glicosilación de Proteínas

Con el objetivo de generar proteínas recombinantes con N-glicanos remodelados, se seleccionaron como dianas la alfa 1-3 fucosiltransferasa y la β 1-2 xilosiltransferasa para el silenciamiento génico de ARNi en *L. minor*. Los resultados iniciales a partir de los esfuerzos de secuenciamiento de ADNc indicaron que había presentes dos o más isoformas para cada uno de los genes diana. Se determinó que la homología de secuencia entre las isoformas estaba entre el 90% y el 95%. Se obtuvieron las secuencias de ADNc de longitud completa correspondientes a ambos genes y se caracterizaron. La secuencia de ADNc de longitud completa, que incluye 5'- y 3'-UTR, para la α 1-3 fucosiltransferasa (FucT) de *L. minor* se establece en la Figura 1; véase también la SEC ID N°: 1 (marco de lectura abierto establecido en la SEC ID N°: 2). La secuencia de aminoácidos predicha codificada por la misma se establece en la SEC ID N°: 3. La proteína codificada comparte alguna similitud con otras FucTs de otras plantas superiores. Véase la Figura 2. Por ejemplo, la secuencia de FucT de *L. minor* comparte aproximadamente un 50,1% de identidad de secuencia con la FucT de *Arabidopsis thaliana* mostrada en la Figura 2.

La secuencia de ADNc de longitud completa, que incluye la 5'- y 3'-UTR, correspondiente a β 1-2 xilosiltransferasa (XylT) de *L. minor* (isoforma n°1) se establece en la Figura 3; véase también la SEC ID N°: 4 (ORF establecido en la SEC ID N°: 5). La secuencia de aminoácidos predicha codificada con ella se establece en la SEC ID N°: 6. La proteína codificada comparte algo de similitud con otras XylTs de otras plantas superiores. Véase la Figura 4. Por ejemplo, la XylT de *L. minor* comparte aproximadamente un 56,4% de identidad de secuencia con la XylT de *Arabidopsis thaliana* mostrada en la Figura 4. Una secuencia de ADNc de longitud parcial, que incluye 3'-UTR, correspondiente a β 1-2 xilosiltransferasa (XylT) de *L. minor* (isoforma n°2) se establece en la Figura 31; véase también la SEC ID N°: 19 (ORF establecido en la SEC ID N°: 20). La secuencia de aminoácidos predicha codificada por la misma se establece en la SEC ID N°: 21. La isoforma n°2 de XylT de longitud parcial comparte una elevada identidad de secuencia con la correspondiente región de la isoforma n°1 de XylT de longitud completa, como puede observarse a partir del alineamiento mostrado en la Figura 32.

Ejemplo 2: Inhibición por ARNi de la Expresión de FucT y XylT de *L. minor*

Se consideraron varias estrategias de ARNi para inhibir la expresión de las isoformas de FucT y XylT de *L. minor*. Las Figuras 5-7, 33 y 34 describen dichas estrategias. Las Figuras 8-13 muestran mapas de las diversas construcciones que se hicieron para lograr el bloqueo deseado de la expresión de estos dos genes. Se generó una

serie de líneas transgénicas que comprenden las diversas construcciones de ARNi de bloqueo usando protocolos de transformación estándares descritos anteriormente en la presente memoria.

El anticuerpo de ensayo, designado en la presente memoria mAbI, se expresó en *Lemna* natural que tiene la maquinaria de glicosilación nativa, y las líneas transgénicas de *Lemna* que expresan construcciones de ARNi diseñadas para inhibir la expresión de isoformas de XylIT y FucT de *L. minor*. Generalmente, se construyeron tres vectores binarios para la expresión de mAbI en el sistema de *Lemna*. El vector de expresión mAbI01 contenía genes de codón optimizado que codifican cadenas pesadas (H) y ligeras (L) de mAbI; el vector mAbI04 contenía genes de codón optimizado que codifican cadenas H y L de mAbI y una construcción de ARNi quimérico dirigida a la expresión de ambas isoformas XylIT y FucT; y el vector mAbI05 contenía genes de codón optimizado que codifican cadenas H y L de mAbI, una construcción de ARNi de gen individual dirigida a la expresión génica de FucT, y una construcción de ARNi de gen individual dirigida a la expresión génica de XylIT. Se generaron líneas transgénicas independientes para los vectores de expresión mAbI01, mAbI04 y mAbI05.

Los genes optimizados para las cadenas H y L de mAbI se diseñaron para tener el uso de codón preferido de *Lemna* (63%-67% de contenido en GC) y contenían la secuencia señal de α -amilasa de arroz (GenBank M24286) fusionada al extremo 5' de sus secuencias codificadoras. Los sitios de restricción de endonucleasa se añadieron para clonar en vectores binarios de *Agrobacterium* (EcoRI (5') / SacI (3'), cadena H) y (Sall (5') / HindIII (3'), cadena L).

Para los datos de XF02 presentados en las Figuras 15-17 y los datos de mAbI04 presentados en las Figuras 22-24 y 26, descritas más adelante en la presente memoria, la estrategia de ARNi para inhibir la expresión de las isoformas de FucT y XylIT de *L. minor* empleó el diseño de ARNi quimérico mostrado en la Figura 34. Para los datos de mAbI05 presentados en la Figura 27, descrita más adelante en la presente memoria, la estrategia de ARNi para inhibir la expresión de las isoformas de FucT y XylIT de *L. minor* empleó un bloque doble de dichos genes usando una combinación de los diseños de ARNi de gen individual mostrados en la Figura 5 (diseño de ARNi de FucT) y la Figura 33 (diseño de ARNi de XylIT).

Se crearon casetes de expresión independientes que contienen promotor, gen de interés y terminador Nos para las cadenas H y L de mAbI optimizado y el ARNi de gen individual o quimérico. Las casetes de expresión fueron clonadas en una modificación del vector binario pBMSP3 de *Agrobacterium* (obtenido del Dr. Stan Gelvin, Universidad de Purdue) con los sitios de restricción apropiados. Dependiendo de la casete de expresión, la cadena L se fusionó con el promotor quimérico de octopina y manopina sintasa con líder 5'RbcS de *Lemna gibba* (mAbI01, Figura 14) o el promotor de poliubiquitina de *Lemna minor* constitutivo de elevada expresión (LmUbq) (mAbI04, Figura 12; mAbI05; Figura 13). La cadena H se fusionó con el promotor quimérico modificado de octopina y manopina sintasa con líder 5' RbcS de *Lemna gibba* (mAbI04, mAbI05 y mAbI01). La casete de ARNi quimérico, tomada del plásmido XF02 de T7-4, se fusionó con el promotor de poliubiquitina de *Spirodela polyrhiza* constitutivo de elevada expresión (SpUbq). La casete de ARNi de gen individual para la expresión de la secuencia inhibidora de FucT fue dirigida por el promotor SpUbq; y la casete de ARNi de gen individual para la expresión de la secuencia inhibidora de XylIT fue dirigida por un elemento de control de la expresión ligado operativamente que comprende el promotor de ubiquitina de *Lemna aequinoctialis* más 5' UTR (promotor LaUbq). Las casetes de expresión de ARNi quimérico fueron clonadas en el vector binario pBMSP3 modificado en orientación tándem creando el plásmido mAbI04. Las casetes de expresión H, L y de ARNi de gen individual dirigidas a la expresión de FucT y XylIT fueron clonadas en el vector binario pBMSP3 modificado creando el plásmido mAbI05. Las casetes de expresión H y L fueron clonadas en el vector binario pBMSP3 modificado creando el plásmido mAbI01.

Aunque se puede usar cualquier protocolo de transformación como se ha indicado en la presente memoria anteriormente, en algunas realizaciones, el protocolo de transformación fue el siguiente. Usando *Agrobacterium tumefaciens* C58Z707, una cepa C58 de rango amplio de hospedante desarmada (Hepburn et al. (1985) J. Gen. Microbiol. 131: 2961-2969), se generaron plantas transgénicas que representan líneas de clones individuales a partir de nódulos de *Lemna minor* de crecimiento rápido según el procedimiento de Yamamoto et al. (2001) In Vitro Cell Dev. Biol. Plant 37: 349-353. Para el escrutinio transgénico, se pre-acondicionaron las líneas clonales individuales durante 1 semana a entre 150 y 200 $\mu\text{mol m}^{-2}\text{s}^{-2}$ en recipientes de crecimiento de plantas venteados que contenían medio SH (Schenk y Hildebrandt (1972) Can. J. Botany 50: 199-204) sin sacarosa. Entre quince y veinte frondas preacondicionadas fueron colocadas entonces en los recipientes ventilados que contenían medio SH fresco, y se dejó que crecieran durante dos semanas. Se congelaron muestras de tejido y medio de cada línea y se almacenaron a -70°C .

Se desarrolló un ensayo MALDI-TOF para medir las actividades de β -1,2-xilosiltransferasa (XylIT) y α -1,3-fucosiltransferasa (FucT) de *L. minor* (Ejemplo 3 mostrado más adelante).

Las Figuras 15-17 representan datos primarios de escrutinio para las líneas vegetales XF02, mAbI04 y mAbI05 obtenidos usando el ensayo mencionado anteriormente. En dicho ensayo, WT (tipo natural) representa la actividad de FucT y XylIT en plantas naturales, mientras que BWT (tipo natural hervido) representa su actividad en extractos de plantas hervidas. Los extractos de planta natural hervida (BWT) son representativos de material de la planta en el que la actividad de FucT y XylIT ha sido desactivada. Este conjunto de datos muestra que varias líneas vegetales del material de la planta tienen un nivel reducido de actividad de FucT y XylIT en comparación con las líneas vegetales naturales (WT), y un nivel de actividad comparable al de las muestras naturales hervidas (BWT).

Específicamente, los datos primarios del escrutinio para líneas vegetales de *L. minor* de ARNi transgénico que comprenden la construcción XF02 de la Figura 10 se muestran en las Figuras 15 y 16. La construcción XF02 expresa una molécula de ARNi quimérico que apunta a la expresión de proteínas FucT y XylT de *L. minor*, incluyendo las varias isoformas de las respectivas proteínas.

- 5 La Figura 17 muestra los datos de escrutinio primarios para las líneas vegetales de *L. minor* de ARNi transgénico que comprenden la construcción mAbl04 de la Figura 12 y la construcción mAbl05 de la Figura 13.

Ejemplo 3: Ensayo MALDI-TOF para Actividad de β -1,2-xilosiltransferasa (XylT) y α -1,3-fucosiltransferasa (FucT) de N-glicano

- 10 Se usó el siguiente ensayo MALDI-TOF modificado para determinar la actividad de XylT y FucT en las plantas transgénicas descritas en el Ejemplo 2 anterior.

Materiales:

TAMPÓN DE HOMOGENEIZACIÓN: HEPES 50 mM, pH 7,5, sacarosa 0,25 M, EDTA 2 mM, DTT 1 mM.

TAMPÓN DE REACCIÓN: Mes 0,1 M, pH 7,0, MnCl_2 10 mM, 0,1% (v/v) de Triton X-100.

URIDINA-5'-DIFOSFO-D-XILOSA (UDP-Xyl)

- 15 GUANOSINA-5'-DIFOSFO-L-FUCOSA (GDP-Fuc)

N-ACETILGLUCOSAMINA

MEZCLA DE POLIETILÉN GLICOL (PEG) 1000-3000 (10 mg/mL de PEG 1000, 2000 y 3000 (relación 4:5:6) mezclado 4:1 con 2 mg/mL de ioduro de sodio).

[Glu¹]-FIBRINOPÉPTIDO B (GFP), HUMANO (1 pmol/mL en agua)

- 20 ACEPTOR DE N-GLICANO DE TETRAPÉPTIDO DABSILADO (EMD Biosciences)

MATRIZ CHCA (ÁCIDO α -CIANO-4-HIDROXICINÁMICO) (10 mg en 50% [v/v] de acetonitrilo, 0,05% [v/v] de ácido trifluoroacético).

Preparación de microsoma:

- 25 Se molió tejido de *L. minor* (100 mg) en 1 mL de tampón de homogeneización frío en una molino de bolas a una velocidad de 5x durante 40 s. El homogenato se giró a 1.000 g durante 5 minutos, 4°C. El sobrenadante se eliminó y se giró a 18.000 g durante 2 h, 4°C. El sobrenadante se desechó de nuevo. La partícula se resuspendió en 20 μL de tampón de reacción frío y se mantuvo en hielo o almacenado a -80°C hasta su uso.

Condiciones de reacción:

- 30 La mezcla de reacción contiene N-acetilglucosamina 125 mM, UDP-Xyl 6,25 mM, GDP-Fuc 6,25 mM, MnCl_2 12,5 mM y 1,5 nmol de aceptor de N-glicano de tetrapéptido dabsilado. Los microsomas (4 μL) se añadieron a la mezcla de reacción para iniciar la reacción. La reacción se incubó durante 30 minutos a temperatura ambiente, y 90 minutos a 37°C. La reacción se terminó mediante centrifugación a 18.000 g durante 1 minuto e incubación a 4°C.

Análisis MALDI-TOF:

- 35 Se mezcló una porción del sobrenadante de cada reacción (0,5 μL) con 0,5 μL de matriz CHCA en una placa diana MALDI y se dejó secar. El instrumento MALDI se fijó en el modo de ion positivo de reflectrón y se calibró con PEG 1000-3000. Se adquirieron los espectros MS combinados (~200 tomas) de 1500-2500 Da usando 0,5 pmol de GFP como masa de cierre. El conteo de iones del pico de referencia ($m/z = 2222,865$) debería estar por encima de 400. El conteo de iones de los productos de XylT y FucT ($m/z = 2192,854$ y $2206,870$, respectivamente) se normalizó al pico de referencia y a la concentración de proteína de la fracción de microsoma.

- 40 Ejemplo 4: Efecto de Inhibición de ARNi de la Expresión de FucT y XylT de *L. minor* sobre el Perfil de Glicosilación de Anticuerpos Monoclonales

- 45 Los anticuerpos monoclonales producidos por *L. minor* natural (es decir, con expresión de FucT y XylT no silenciada) que comprenden la construcción mAbl01 (véase la Figura 14) y las líneas de *L. minor* transgénica para la construcción mAbl04 (véase la Figura 12) o la construcción mAbl05 (véase la Figura 13) se analizaron para determinar su perfil de N-glicosilación. Se usaron los siguientes procedimientos.

Purificación de mAb de *Lemna*:

Se homogeneizó tejido vegetal con Fosfato Sódico 50 mM, Cloruro Sódico 0,3M y EDTA 10 mM a pH 7,2 usando un Mezclador de Alta Cizalla Silverson con una relación de tejido:tampón de 1:8. El homogenato se acidificó a pH 4,5 con Ácido Cítrico 1 M, y se centrifugó a 7.500 x g durante 30 minutos a 4°C. Se filtro el sobrenadante a través de un filtro de 0,22 µm y se cargó directamente en resina SuRe mAbSelect (GE Healthcare) equilibrada con una disolución que contiene Fosfato Sódico 50 mM, Cloruro Sódico 0,3M y EDTA 10 mM, pH 7.2. Tras cargar, la columna se lavó hasta línea base con el tampón de equilibrado seguido de un lavado intermedio con 5 volúmenes de columna de Acetato Sódico 0,1M, pH 5,0 y finalmente, se eluyó el anticuerpo ligado con 10 volúmenes de columna de Acetato Sódico 0,1 M, pH 3,0. El eluato se neutralizó inmediatamente con base Tris 2M.

Purificación de glicanos *N*-ligados:

Los anticuerpos monoclonales purificados de proteína A (1 mg) procedentes de líneas vegetales de *L. minor* de ARNi y natural fueron dializados extensivamente contra agua y liofilizados hasta sequedad. Las muestras fueron resuspendidas en 100 µL de ácido fórmico al 5% (v/v), llevadas a 0,05 mg/mL de pepsina e incubadas a 37°C durante la noche. Las muestras fueron desactivadas térmicamente a 95°C durante 10 minutos y se secaron. Los digestos de pepsina fueron resuspendidos en 100 µL de acetato sódico 100 mM, pH 5,0 y se incubaron con 1 mU de *N*-glicosidasa A a 37°C durante la noche. Los *N*-glicanos liberados se aislaron usando 4 cm³ de columnas Carbograph SPE según Packer et al. (1998) Glycoconj. J. 15: 737-747, y se secaron.

Los *N*-glicanos secos fueron purificados adicionalmente usando 1 cm³ de cartuchos Waters Oasis MCX. Las columnas se prepararon lavando con 3 volúmenes de columna de metanol seguido de 3 volúmenes de columna de ácido fórmico al 5% (v/v). Los *N*-glicanos, resuspendidos en 1 mL de ácido fórmico al 5% (v/v), fueron cargados en las columnas preparadas. La fracción no ligada, así como 2 volúmenes de lavado columna adicionales de ácido fórmico al 5% (v/v), se reunieron y se secaron.

Derivatización de oligosacáridos con ácido 2-aminobenzoico (2-AA):

Los *N*-glicanos o maltooligosacáridos purificados se marcaron con 2-AA y se purificaron usando cartuchos Waters Oasis HLB de 1 cm³ según Anumula y Dhume (1998) Glycobiology 8: 685-694. Los *N*-glicanos y maltooligosacáridos marcados se resuspendieron en 50 µL de agua y se analizaron mediante MALDI-TOF MS y HPLC-QTOF MS de fase normal (NP).

Espectrometría de Masas MALDI-TOF:

Se llevó a cabo la MALDI-TOF MS usando un Waters MALDI Micro MX (Milford, MA). Se diluyeron *N*-glicanos marcados con 2-AA (0,5 µL) apropiadamente con agua, se mezclaron con 0,5 µL de matriz DHB 10 mg/mL en acetonitrilo al 70% (v/v), se visualizó sobre una placa diana y se analizó en modo de reflectrón negativo.

Análisis NP-HPLC-Q-TOF MS de *N*-glicanos marcados con 2-AA:

Los *N*-glicanos o maltooligosacáridos marcados con 2-AA se llevaron a acetonitrilo al 80% (v/v) y se separaron en un sistema de HPCL Waters 2695 ajustado con una columna TSK-Gel Amide-80 (2 mm x 25 cm, 5 µm) (Tosoh Biosciences, Montgomeryville, PA). Los carbohidratos marcados con 2-AA se detectaron y se analizaron mediante fluorescencia (excitación a 230 nm, emisión a 425 nm) usando un detector de fluorescencia Waters 2475 y un espectrómetro de masas de cuadrupolo-tiempo de vuelo (Q-TOF) Waters Q-TOF API US (Millford, MA) ajustado en línea con el sistema de HPLC.

Las separaciones se llevaron a cabo a 0,2 mL/min, 40°C, usando acetato de amonio 10 mM, pH 7,3 (disolvente A) y acetato de amonio 10 mM, pH 7,3, acetonitrilo al 80% (v/v) (disolvente B). La elución de la muestra se llevó a cabo con un 0% de A isocrático durante 5 minutos, seguido de un aumento lineal hasta un 10% de A a los 8 minutos, y un aumento lineal hasta un 30% de A durante 15 minutos antes de la siguiente inyección.

El análisis Q-TOF se llevó a cabo en modo de ión negativo con temperaturas de fuente y de-solvatación de 100°C y 300°C, respectivamente, y voltajes capilares y de cono de 2.100 y 30 V, respectivamente. Los espectros de masas mostrados son el resultado de combinar ≥50 escaneos individuales por *N*-glicano marcado.

Análisis RP-HPLC-Q-TOF MS de IgG intacta:

IgG's purificadas de proteína A (50 µg) se desalinizaron usando el sistema de HPLC Waters 2695 ajustado a una columna Poros R1-10 (2 mm x 30 mm; Applied Biosystems). Las IgG's fueron detectadas y analizadas usando un detector UV de longitud de onda dual Waters 2487 (280 nm) y el Waters Q-TOF API US. Las separaciones se llevaron a cabo a 0,15 mL/min, 60°C, usando ácido trifluoroacético al 0,05% (v/v) (TFA; disolvente A) y TFA al 0,05% (v/v), acetonitrilo al 80% (v/v) (disolvente B). La elución de la muestra se llevó a cabo usando un incremento lineal desde 30 hasta 50% de B durante 5 minutos, un incremento hasta 80% de B durante 5 minutos. La relación de disolventes se mantuvo a 80% de B durante otros 4 minutos, seguido de un lavado con 100% de B durante 1 minutos y un equilibrado de la columna con un 30% de B durante 15 minutos antes del siguiente experimento.

Se llevó a cabo el análisis Q-TOF en modo de ión positivo con temperaturas de fuente y de de-solvatación de 100°C y 300°C, respectivamente, y voltajes capilares y de cono de 3,0 y 60 V, respectivamente. Los datos son el resultado de combinar ≥ 100 escaneos individuales y de una deconvolución respecto al espectro de masas original usando MaxEnt 1.

- 5 Véase también Triguero et al. (2005) Plant Biotechnol. J. 3: 449-457; Takahashi et al. (1998) Anal. Biochem. 255: 183-187; Dillon et al. (2004) J. Chromatogr. A. 1053: 299-305.

Resultados:

La Figura 18 muestra la estructura y el peso molecular de N-glicanos de anticuerpo monoclonal de *L. minor* natural derivatizado.

- 10 La Figura 19 muestra que la construcción mAbI01 natural (mostrada en la Figura 15) que proporciona la expresión del anticuerpo de IgG 1 monoclonal mAbI en *L. minor*, sin supresión de ARNi de FucT y XylIT de *L. minor*, produce un perfil de N-glicosilación con tres especies de N-glicano principales, que incluyen una especie que tiene residuos de xilosa β 1,2-ligada y una especie que tiene residuos de xilosa β 1,2-ligada y fucosa α 1,3-ligada de núcleo; dicho perfil se confirma mediante análisis de espectrometría de masas de cromatografía de líquidos (LC-MS) (Figura 20) y MALDI (Figura 21).

- 15 La Figura 22 muestra un solapamiento de las cantidades relativas de las diversas especies de N-glicano del mAbI producido en la línea de *L. minor* natural que comprende la construcción de mAbI0 1 (sin supresión de FucT o XylIT) y en las dos líneas de *L. minor* transgénica que comprenden la construcción mAbI04 de la Figura 12. Nótese el enriquecimiento de la especie de glicano GnGn (es decir, G0), sin residuos unidos de xilosa β 1,2-ligada o de fucosa α 1,3-ligada de núcleo, y la ausencia de especies que tienen residuos de xilosa β 1,2-ligada o residuos de xilosa β 1,2-ligada y fucosa α 1,3-ligada de núcleo. Dicho perfil se confirma con análisis de espectrometría de masas (LC-MS) (Figura 23) y MALDI (Figura 24).

- 20 La Figura 25 muestra el análisis de masas intactas de las composiciones mAbI producidas en *L. minor* natural (línea 20) que comprende la construcción mAbI01. Cuando la expresión de XylIT y FucT no es suprimida en *L. minor*, la composición de mAbI producida recombinantemente es heterogénea, comprendiendo al menos 9 glicoformas diferentes, siendo la glicoforma GOXF³ la especie predominante de las presentes. Nótese el pico muy pequeño que representa a la glicoforma G0.

- 25 La Figura 26 muestra el análisis de masas intactas de las composiciones de mAbI producidas en *L. minor* transgénicas (línea 15) que comprende la construcción mAbI04 de la Figura 12. Cuando las expresiones de XylIT y FucT son suprimidas en *L. minor* usando esta construcción de ARNi quimérico, la composición de mAbI intacta es sustancialmente homogénea para N-glicanos G0, con tan solo cantidades traza de N-glicanos precursores (representados por las especies de glicano precursoras GnM y MGn). Adicionalmente, la composición de mAbI es sustancialmente homogénea para la glicoforma G0, en donde ambos sitios de glicosilación están ocupados por las especies de N-glicano G0, con tres picos menores que reflejan cantidades traza de glicoformas precursoras (un pico que muestra mAbI que tiene una región Fc en la que el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de la otra cadena pesada está sin glicosilar; otro pico que muestra mAbI que tiene una región Fc en la que el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unida a Asn 297, y el dominio C_H2 de la otra cadena pesada tiene el glicano precursor GnM o MGn unido a Asn 297; y otro pico que muestra mAbI que tiene una región Fc en donde el sitio de glicosilación de Asn 297 de cada uno de los dominios C_H2 tiene la especie de glicano G0 unida, con una tercera especie de glicano G0 unida a un sitio de glicosilación adicional de la estructura de mAbI).

- 30 La Figura 27 muestra el análisis de masas intactas de las composiciones de mAbI producidas en *L. minor* transgénica (línea 72) que comprende la construcción mAbI05 de la Figura 13. Cuando las expresiones de XylIT y FucT son suprimidas en *L. minor* usando esta construcción, la composición mAbI intacta es sustancialmente homogénea para N-glicanos G0, con tan solo cantidades traza de especies de N-glicano precursoras (representadas por las especies de glicano precursoras GnM y MGn). Adicionalmente, la composición de mAbI es sustancialmente homogénea (al menos un 90%) para la glicoforma G0, con los mismos tres picos menores que reflejan glicoformas precursoras obtenidos con la construcción mAbI04.

- 35 La actividad de unión de receptor del mAbI producido en las líneas de *Lemna* naturales que comprenden la construcción mAbI01 (es decir, sin inhibición de la expresión de XylIT y FucT) y las líneas de *Lemna* transgénica que comprenden las construcciones mAbI04 y mAbI05 (es decir, con las expresiones de XylIT y FucT inhibidas) se comparó con la actividad de unión de receptor de mAbI producido en líneas celulares de mamífero (CHO y SP2/0).

- 40 Se determinó la unión a FcFcγRIIIa en células NK humanas recién aisladas para los diversos productos mAbI. Los datos de control recogidos para mAbI derivado de CHO y mAbI derivado de SP2/0 se muestran en la Figura 35. Los datos de ensayo recogidos para mAbI producido en *Lemna* natural que tiene el perfil de N-glicanos de planta normal se designan como mAbI01-15 y mAbI01-20, en donde el producto mAbI tiene glicanos N-ligados que incluyen residuos de α (1,3)-fucosa (véanse las Figuras 36 y 37). Los datos de ensayo recogidos para mAbI producido en *Lemna* transgénica que tiene un perfil de N-glicanos optimizado (OPT) obtenida con construcciones de ARNi

silenciadoras de genes que apuntan a la expresión de $\alpha(1,3)$ -fucosiltransferasa se designan como mAbI05-72, mAbI05-74, mAbI04-24 y mAbI04-15 (véanse las Figuras 36 y 37), en donde el producto mAbI tiene glicanos N-ligados que están desprovistos de residuos de $\alpha(1,3)$ -fucosa. En la Figura 38 se muestran los datos que comparan la eficacia de unión de mAbI01-15, mAbI01-20, mAbI SP2/0, mAbI04-15, mAbI04-24, mAbI05-72 y mAbI05-74 a Fc γ RIV recombinante de ratón, un receptor que es sensible a los niveles de fucosa de IgG y que sirve como sustituto de la Fc γ RIIIa humana.

Estos datos demuestran que el producto mAbI derivado de *Lemna* transgénica que tiene el perfil de glicanos optimizado (OPT) muestra una mayor unión a Fc γ RIIIa en células NK humanas recién aisladas (potenciada entre aproximadamente 20 y 50 veces) así como una mayor unión a Fc γ RIV recombinante de ratón (potenciada aproximadamente 10 veces) en comparación con el producto mAbI derivado de *Lemna* natural.

Ejemplo 6: Producción de Anticuerpo Monoclonal Anti-CD30 que Tiene una Unión a Receptor Mejorada y una Actividad ADCC Incrementada

Este ejemplo describe la expresión de mAbs anti-CD30 humanos en *Lemna*. La optimización de la glicosilación de mAb anti-CD30 se llevó a cabo mediante la co-expresión con una construcción de ARNi dirigida a la expresión endógena de genes de α -1,3-fucosiltransferasa (FucT) y β -1,2-xilosiltransferasa (XylT) de un modo similar al indicado en los anteriores ejemplos para mAbI. El mAb anti-CD30 resultante producido en *Lemna* a la que se ha modificado su maquinaria de glicosilación nativa para suprimir la expresión de FucT y XylT contenía una única especie de N-glicano principal sin ninguna traza de N-glicanos específicos de planta. Además de la homogeneidad de N-glicanos, los mAbs anti-CD30 glico-optimizados también demostraron tener una citotoxicidad mediada por célula dependiente de anticuerpo (ADCC) potenciada, y una actividad de unión a receptor de célula efectora potenciada, cuando se compararon con mAbs anti-CD30 expresados en CHO.

Métodos

Cepas y Reactivos:

Se usaron células de *Escherichia coli* competentes Novablu para todo el trabajo de ADN recombinante (EMD Biosciences, San Diego, CA). Se obtuvieron endonucleasas de restricción y enzimas de modificación de ADN a partir de New England Biolabs (Ipswich, MA). Los oligonucleótidos se obtuvieron de Integrated DNA Technologies (Coralville, IA). Las columnas Waters Oasis HLB y MCX (1 cm³), el ácido 2,5-dihidroxibenzoico (DHB) y el ácido α -ciano-4-hidroxicinámico (CHCA) provenían de Waters Corporation (Milford, MA). Los aceptores de N-glicano GnGn de tetrapéptido dabsilado (GnGn-dabsil-péptido) y la N-glicosidasa A provenían de EMD Biosciences. Las columnas Carbohydrate SPE (4 cm³) provenían de Grace Davidson Discovery Sciences (Deerfield, IL). La uridina-5'-difosfo-D-xilosa (UDP-Xyl) se adquirió en Carbosource Services (Athens, GA). El acetonitrilo (grado Optima) provenía de Fisher Scientific (Summerville, NY). El acetato de amonio era de MP Biochemicals (Irvine, CA). Los maltooligosacáridos (MD6-1) provenían de V-Labs Inc. (Covington, CA). Los patrones de monosacárido provenían de Dionex (Sunnyvale, CA). BATDA (bis(acetoximetil)2,2':6',2''-terpiridina-6,6''-dicarboxilato) y la disolución de Europio provenían de Perkin-Elmer (Wellesley, MA). La guanosina-5'-difosfo-L-fucosa (GDP-Fuc), la N-acetilglucosamina (GlcNAc), el ácido 2-aminobenzoico (2-AA) y todos los demás materiales procedían de Sigma (St. Louis, MO).

Construcción de mAb y vectores de expresión de ARNi:

Se determinaron las secuencias de ADNc de región variable de cadena pesada (H) y ligera (L) de anticuerpo kappa mAb1 MDX-060 completamente humano derivado de un Medarex HuMAb-Mouse® transgénico (Borchmann et al. (2003) Blood 102: 3737-3742) y se produjo el anticuerpo kappa mAb1 MDX-060 de longitud completa recombinantemente mediante una línea de células de ovario de hámster chino, CHO DG44 (Urlaub et al. (1986) Cell Mol. Genet. 12: 555-566), usando técnicas estándar. Se diseñaron los genes optimizados correspondientes a las cadenas H y L para tener un uso de codón preferido de *Lemna* (63%-67% de contenido en GC) y contienen la secuencia señal de α -amilasa (GenBank M24286) fusionada al extremo 5' de sus secuencias codificadoras. Los sitios de restricción de endonucleasa se añadieron para clonar en vectores binarios de *Agrobacterium* (EcoRI (5') / SacI (3'), cadena H) y (Sall (5') / HindIII (3'), cadena L). Se construyeron los genes sintéticos y se proporcionaron mediante Picoscript (Houston, TX).

Se diseñó un ARN de horquilla quimérico (véase la Figura 34) dirigido a silenciar genes endógenos de *Lemna* que codifican α -1,3-fucosiltransferasa (en base a la secuencia codificadora correspondiente a la isoforma n°1 de FucT de *L. minor*, establecida en la SEC ID N°: 2; véase también GenBank DQ789145) y β -1,2-xilosiltransferasa (en base a la secuencia codificadora correspondiente a la isoforma n°2 de XylT de *L. minor*, nt 1-1275 de la SEC ID N°: 19; establecida en la SEC ID N°: 20; véase también GenBank DQ789146). El ARN de horquilla de FucT + XylT se diseñó para tener 602 pb de secuencia de FucT de cadena doble, 626 pb de secuencia de XylT de cadena doble, y 500 pb de secuencia espaciadora. La porción de cadena sentido de la casete de ARN de horquilla abarca el fragmento directo de FucT (nt 12-613 de la SEC ID N°: 2; equivalente a los nt 254-855 de la SEC ID N°: 1) y el fragmento directo de XylT (nt 1-626 de la SEC ID N°: 19), una secuencia espaciadora (nt 627-1126 de la SEC ID N°: 19). La porción de cadena antisentido del ARN de horquilla abarca el fragmento inverso de XylT (versión antisentido

de los nt 1-626 de la SEC ID N°: 19) y el fragmento inverso de FucT (versión antisentido de los nt 12-613 de la SEC ID N°: 2 o los nt 254-855 de la SEC ID N°: 1). El ARN de horquilla quimérico se construyó mediante PCR amplificando los fragmentos génicos directos e inversos de FucT y XylT procedentes de ADNc de *Lemna minor* (8627) y clonándolos secuencialmente en pT7blue (EMD Biosciences) creando el plásmido XF02 en T7-4. El fragmento génico directo de FucT se amplificó con los cebadores de ADN BLX 686 (5'-ATGGTCGACTGCTGCTGGTGCTC TCAAC-3') (SEC ID N°: 22) y BLX690 (5'-ATGTCTAGAATG CAGCAGCAAGT GCACC-3') (SEC ID N°: 23) produciendo un producto de 620 pb con sitios de clonación terminales Sall (5') y XbaI (3'). El fragmento génico directo de XylT se amplificó con los cebadores de ADN BLX 700 (5'-ATGACTAGTTGC GAAGCCTACTTCGGCAACAGC3') (SEC ID N°: 24) y BLX694 (5'-ATGGGATCCGAATCTCAAGA ACAACTGTGCG-3') (SEC ID N°: 25) produciendo un producto de 1144 pb con sitios de clonación terminales SpeI (5') y BamHI (3'). El fragmento génico inverso de XylT se amplificó con los cebadores de ADN BLX 695 (5'-ATGGGTACCTGCGAAGCCTACTTCGGCAACAGC-3') (SEC ID N°: 26) y BLX696 (5'-ATGGGATCCACTGGCTGG GAGAAGTTCTT-3') (SEC ID N°: 27) produciendo un producto de 644 pb con sitios de clonación terminales BamHI (5') y KpnI (3'). El fragmento génico inverso de FucT se amplificó con los cebadores de ADN BLX 691 (5'-ATGGAGCTCTGCTGCTGGTGCTCTCAAC-3') (SEC ID N°: 28) y BLX692 (5'-ATGGGTACCATGCAGCAGCAA GTGCACC-3') (SEC ID N°: 29) produciendo un producto de 620 pb con sitios de clonación terminales KpnI (5') y SacI (3').

Se crearon casetes de expresión independientes que contienen promotor, el gen de interés y terminador Nos para las cadenas H y L de MDX-060 optimizado y para el ARNi quimérico. Las casetes de expresión se clonaron en una modificación del vector binario de *Agrobacterium* pBMSP3 (obtenido del Dr. Stan Gelvin, Universidad de Purdue) con los sitios de restricción apropiados. La cadena H se fusionó al promotor de octopina y manopina sintasa quimérico con líder 5' RbcS de *Lemna gibba* (Gasdaka et al. (2003) Bioprocessing J. 50-56). La cadena L se fusionó con el promotor de poliubiquitina de *Lemna minor* constitutivo de elevada expresión (LmUbq). La casete de ARNi quimérico, tomada del plásmido XF02 de T7-4, se fusionó con el promotor de poliubiquitina de *Spirodela polyrrhiza* constitutivo de elevada expresión (SpUbq). Las tres casetes de expresión fueron clonadas en el vector binario pBMSP3 modificado en orientación tándem creando el plásmido MDXA04.

Transformación y escrutinio de línea vegetal:

Usando *Agrobacterium tumefaciens* C58Z707 (Hepburn et al. (1985) J. Gen. Microbiol. 131: 2961-2969), se generaron plantas transgénicas que representan líneas clonales individuales a partir de nódulos de *Lemna minor* de crecimiento rápido según el procedimiento de Yamamoto et al. (Yamamoto et al. (2001) In vitro. Cell. Dev. Biol. 37). Para el escrutinio transgénico, se pre-acondicionaron las líneas clonales individuales durante 1 semana a entre 150 y 200 $\mu\text{mol m}^{-2}\text{s}^{-2}$ en recipientes de crecimiento de plantas ventilados que contenían medio SH (Schenk y Hildenbrandt (1972) Can. J. Botany 50: 199-204) sin sacarosa. A continuación se colocaron de quince a veinte frondas preacondicionadas en los recipientes ventilados que contenían medio SH fresco y se dejaron crecer durante dos semanas. Se congelaron muestras de tejido y de medio para cada línea y se almacenaron a -70°C.

Análisis ELISA de mA b producidos en *Lemna*:

Se homogeneizó tejido de *Lemna* (100 mg) usando un molino de bolas FastPrep FP120 (Thermo Electron Corporation). Los sobrenadantes se diluyeron hasta 1 $\mu\text{g/mL}$ y se ensayaron usando el kit ELISA de cuantificación de IgG (Bethyl Laboratories). Para el ensayo, las placas microtiter se recubrieron con IgG anti-humana de cabra a una concentración de 10 $\mu\text{g/mL}$, se detectaron los mA b mediante IgG anti-humana de cabra conjugada con peroxidasa de rábano (HRP) diluido 1: 100.000. Se crearon curvas de calibración con IgG de Referencia Humana suministrada con el kit ELISA. La sensibilidad del ELISA fue de 7,8 ng/mL. Todas las muestras se analizaron por duplicado.

Preparación de membranas microsomales de *Lemna* y determinación de las actividades de β -1,2-xilosiltransferasa y α -1,3-fucosiltransferasa de núcleo:

Se homogeneizó tejido de *Lemna* (100 mg) de cada línea en 1 mL de tampón de homogeneización frío (ácido 4-[2-hidroxietil]-1-piperacinetanosulfónico [HEPES] 50 mM, pH 7,5, sacarosa 0,25 M, ácido etilendiamintetraacético [EDTA] 2 mM, 1,4-ditiotreitol [DTT] 1 mM) durante 40 s en un molino de bolas FastPrep FP120 (Thermo Electron Corporation, Waltham, MA). Se centrifugó el homogenato a 1.000 g durante 5 minutos a 4°C. Se retiró el sobrenadante y se centrifugó a 18.000 g durante 90 minutos a 4°C. La partícula resultante se resuspendió en 20 μL de tampón de reacción frío (ácido 2-[4-morfolino]-etanosulfónico [Mes] 0,1 M, pH 7,0, 0,1% [v/v] de Triton X-100, MnCl_2) y se mantuvo en hielo o almacenado a -80°C hasta su uso.

Se midieron las actividades de β -1,2-xilosiltransferasa y α -1,3-fucosiltransferasa de núcleo simultáneamente en 4 μL de membranas microsomales preparadas a partir de cada línea de ARNi mediante incubación con GlcNAc 125 mM, UDP-Xyl 6,25 mM, GDP-Fuc 6,25 mM, MnCl_2 12,5 mM y 1,5 nmol de aceptor de GnGn-dabsil-péptido durante 2 h a 37°C tal como se ha descrito previamente (Leiter et al. (1999) J. Biol. Chem. 274: 21830-21839). Se finalizó la reacción mediante una breve centrifugación e incubación a 4°C y los productos fueron analizados mediante MALDI-TOF MS en modo reflectrón positivo.

Purificación de mAbs MDX-060 LEX y LEX^{Opt}:

- Se homogeneizó tejido vegetal con fosfato sódico 50 mM, cloruro sódico 0,3 M y EDTA 10 mM, pH 7,2, usando un mezclador de alta cizalla Silverson. El homogenato se acidificó hasta pH 4,5 con ácido cítrico 1 M, y se centrifugó a 7.500 g durante 30 minutos a 4°C. Se ajustó el pH del sobrenadante a 7,2 con Tris 2 M, antes de filtrarse usando filtros de 0,22 µm. El material se cargó directamente en una resina de proteína A mAbSelect SuRe (GE Healthcare) equilibrada con una disolución que contenía fosfato sódico 50 mM, cloruro sódico 0,3 M y EDTA 10 mM, pH 7,2. Tras cargar, la columna se lavó hasta línea base con el tampón de equilibrado seguido de un lavado intermedio con 5 volúmenes de columna de acetato sódico 0,1 M, pH 5,0. Los anticuerpos ligados se eluyeron con 10 volúmenes de columna de acetato sódico 0,1 M, pH 3,0. El eluato de proteína A se neutralizó inmediatamente con 2-amino-2-[hidroximetil]-1,3-propanodiol (Tris) 2 M. Para la eliminación de agregados, el eluato A se ajustó a pH 5,5 y se aplicó a una columna de hidroxiapatita cerámica de tipo I (Bio-Rad) equilibrada con fosfato sódico 25 mM, pH 5,5. Tras lavar la columna con 5 volúmenes de columna de tampón de equilibrado, el anticuerpo se eluyó en una única etapa de elución usando fosfato sódico 0,25 M, pH 5,5. Las fracciones que contenían anticuerpo fueron reunidas y almacenadas a -80°C.
- Las muestras de extracto de tejido y de flujo de proteína A se prepararon para SDS-PAGE en condiciones reductoras y no reductoras mediante la adición de 2x tampón de muestreo SDS ± 5% [v/v] de 2-mercaptoetanol. Las muestras de eluato de proteína A y de eluato de hidroxiapatita se diluyeron hasta una concentración de proteína de 0,5 mg/mL, seguido de la adición de 2x tampón de muestreo SDS ± 5% [v/v] de 2-mercaptoetanol. Las muestras se incubaron a 95°C durante 2 minutos antes de electroforesis usando geles de gradiente de Tris-Glicina del 4-20% (Invitrogen, Carlsbad, CA). Se incluyeron en los geles marcadores de peso molecular Mark 12 (Invitrogen) y un patrón de referencia MDS-060. Los geles se tiñeron con tinte Azul Coloidal.

Purificación de glicanos N-ligados:

- Los anticuerpos monoclonales purificados de proteína A (1 mg) procedentes de líneas vegetales de *Lemna* natural y de ARNi fueron dializados extensivamente contra agua y se liofilizaron hasta sequedad. Las muestras fueron resuspendidas en 100 µL de ácido fórmico al 5% (v/v), se llevaron a pepsina 0,05 mg/mL y se incubaron a 37°C durante la noche. Las muestras fueron desactivadas térmicamente a 95°C durante 15 minutos y fueron secadas. Los digestos de pepsina se resuspendieron en 100 µL de acetato sódico 100 mM, pH 5,0, y se incubaron con 1 mU de N-glicosidasa A a 37°C durante la noche. Los N-glicanos liberados fueron aislados usando columnas Carbohydrate SPE de 4 cm³ (Packer et al. (1998) Glycoconj. J. 19: 737-747) y fueron secados.
- Los N-glicanos secados fueron purificados adicionalmente usando cartuchos Waters Oasis MCX de 1 cm³. Las columnas se prepararon lavando con 3 volúmenes de columna de metanol seguidos de 3 volúmenes de columna de ácido fórmico al 5% (v/v). Los N-glicanos, resuspendidos en 1 mL de ácido fórmico al 5% (v/v) fueron cargados en las columnas preparadas. La fracción no ligada, así como dos volúmenes de columna adicionales de lavado con ácido fórmico al 5% (v/v), fueron recogidos, reunidos y secados.
- Derivatización de oligosacáridos con ácido 2-aminobenzoico (2-AA).*
- Los N-glicanos o maltooligosacáridos purificados fueron marcados con 2-AA y se purificaron usando cartuchos Waters Oasis HLB de 1 cm³ según Anumula y Dhume, 1998 (Anumula y Dhume (1998) Glycobiology 8: 685-694). Los N-glicanos y maltooligosacáridos se resuspendieron en 50 µL de agua y se analizaron mediante MALDI-TOF MS en modo negativo y mediante NP-HPLC-QTOF MS.

Espectrometría de Masas MALDI-TOF:

- Se llevó a cabo un análisis MALDI-TOF MS usando un equipo Waters MALDI Micro MX (Millford, MA). El análisis de productos de reacción de β-1,2-xiloxiltransferasa/α-1,3-fucosiltransferasa se llevó a cabo mezclando 0,5 µL de cada sobrenadante de reacción con 0,5 µL de CHCA a 10 mg/mL en TFA al 0,05% (v/v), acetonitrilo al 50% (v/v) en una placa diana. Los productos de GnGn-dabsil-péptidos xilosilados ([M+H]⁺ = 2192,85 Da) o fucosilados ([M+H]⁺ = 2206,87 Da) fueron detectados en modo de reflectrón positivo. El conteo de iones de 200 espectros combinados para cada muestra fue normalizado frente a GnGn-dabsil-péptido β-1,4-galactosilado ([M+H]⁺ = 2222,87 Da) presente como contaminante (<5%) en la mezcla original de GnGn-dabsil-péptidos de EMD Biosciences.

- Los N-glicanos o maltooligosacáridos marcados con 2-AA (0,5 µL) fueron diluidos con agua, mezclados con 0,5 µL de matriz DHB a 10 mg/mL en acetonitrilo al 70% (v/v), se llevaron a una placa diana y se analizaron en modo de reflectrón negativo.

Análisis NP-HPLC-Q-TOF MS de N-glicanos marcados con 2-AA:

- Los N-glicanos o maltooligosacáridos marcados con 2-AA fueron llevados a acetonitrilo al 80% (v/v) y separados en un sistema Waters 2695 HPLC equipado con una columna TSK-Gel Amide-80 (2 mm x 25 cm, 5 µm) (Tosoh Biosciences, Montgomeryville, PA). Los carbohidratos marcados con 2-AA fueron detectados y analizados usando un detector de fluorescencia Waters 2475 (excitación a 230 nm, emisión a 425 nm) y un espectrómetro de masas de cuadrupolo-tiempo de vuelo (QTOF) Waters Q-TOF API US ajustado en línea con el sistema HPLC.

Las separaciones se llevaron a cabo a 0,2 mL/min, 40°C, usando acetato de amonio 10 mM, pH 7,3 (disolvente A) y acetato de amonio 10 mM, pH 7,3, acetonitrilo al 80% (v/v) (disolvente B). La elución de muestra se llevó a cabo a 0% de A isocrático durante 5 minutos, seguido de un aumento lineal hasta 10% de A a los 8 minutos, y un aumento lineal hasta 30% de A a los 48 minutos. La columna se lavó con un 100% de A durante 15 minutos y se equilibró con 0% de A durante 15 minutos antes de la siguiente inyección.

El análisis QTOF se llevó a cabo en modo de ion negativo con temperaturas de fuente y de desolvatación de 100°C y 300°C, respectivamente, y con voltajes capilar y de cono de 2.100 y 30 V, respectivamente. Los espectros de masas mostrados son el resultado de combinar ≥ 40 escaneos individuales por *N*-glicano marcado.

Análisis de monosacárido mediante HPAEC-PAD:

Las muestras de mAb (200 µg) fueron sometidas a hidrólisis ácida usando TFA 2 N a 100°C durante 3 h. Las muestras se secaron mediante centrifugación a vacío a temperatura ambiente y se reconstituyeron en 150 µL de agua antes de su análisis con HPAE-PAD (Dionex). Se aplicó una alícuota (25 µL) de la muestra reconstituida a una columna CarboPac PA10 (4 x 250 mm) con una pre-columna Amino Trap (Dionex). La separación de los monosacáridos se logró con una fase móvil de KOH 3,5 mM, usando un generador de eluyente EG40. La identidad de los picos de monosacárido y la abundancia relativa fueron determinadas usando patrones de monosacáridos.

Estabilidad térmica de mAb:

Se usó un instrumento de calorimetría de barrido diferencial (DSC, del inglés "Differential Scanning Calorimetry") VP-Capillary de MicroCal (Northampton, MA) para determinar la estabilidad térmica de mAbs naturales y de glicol optimizado. Las muestras de mAb purificados fueron dializadas en NaH₂PO₄ 20 mM, pH 7,4, NaCl 150 mM (PBS) durante la noche. Los datos de desnaturalización térmica se obtuvieron calentando las muestras a una concentración de 300 µg/mL entre 35 y 95°C con una velocidad de barrido de 1°C/min usando PBS como tampón de referencia. La retroalimentación y la ganancia se fijaron bajas. Los datos normalizados con corrección de línea base se ajustaron a un modelo non-2-state usando el software Origin v7.0.

Actividad de unión FcR de los mAb:

Se llevó a cabo el experimento usando un instrumento BIACORE (Biacore AB, Uppsala, Suecia) que usa tecnología de resonancia de plasmón superficial. Los mAbs, 2 µg/mL, fueron capturados en la superficie recubierta de antígeno (CD30 humano recombinante). Sobre los anticuerpos capturados se hizo circular durante 3 minutos concentraciones diversas de los alo tipos Val¹⁵⁸ y Phe¹⁵⁸ de FcRγIIIa, comenzando desde 6 µM. Se ajustó la señal de unión en función del FcRγIIIa a un modelo de un solo sitio de unión usando el software GraphPad Prism (v4.01) para obtener valores de K_D. En todo el experimento se usó tampón HBS-EP (HEPES 10 mM, NaCl 0,15 M, EDTA 3 mM y P20 al 0,005% (v/v) a pH 7,4). Como control negativo se usó la unión de los mAbs a tampón o de FcRγIIIa a superficies limpias.

Ensayo para determinar la afinidad de unión a antígeno:

Para determinar la unión se usaron como células positivas de antígeno células L540 que expresan CD30 (DSMZ Cell Culture Collection n° ACC 72). Se incubaron alícuotas de 2×10^5 células/pocillo durante 30 minutos a 4°C con 100 µL de anticuerpos primarios a las concentraciones indicadas. Las células fueron lavadas dos veces en PBS con suero fetal bovino (FBS) al 2% (v/v) antes de la adición de mAb anti-humano de cabra, anticuerpo secundario marcado con FITC (Jackson ImmunoResearch, West Grove, PA) a una dilución 1:500 en 100 µL/pocillo durante 30 minutos a 4°C. Las células fueron lavadas dos veces en PBS con FBS al 2% (v/v) y se analizaron mediante citometría de flujo usando un instrumento FACS Calibur (Becton Dickinson, Franklin Lakes, NJ). Se determinaron los valores EC₅₀ de la unión de mAbs MDX-060 CHO, LEX y LEX^{Opt} a CD30 sobre células L540 a partir de las curvas de unión utilizando el software GraphPad Prism 3.0.

Ensayo de ADCC:

Se purificaron células efectoras mononucleares sanguíneas periféricas a partir de sangre entera heparinizada mediante separación estándar Ficoll-Paque. Las células (2×10^6) fueron lavadas en PBS y enviadas a genotipado. A continuación, las células efectoras restantes fueron resuspendidas a 1×10^6 células/mL en medio RPMI 1640 que contenía FBS al 10% (v/v) y 50 U/mL de IL-2 humana (Research Diagnostics, Concord, MA) y se incubaron durante la noche a 37°C. Las células efectoras fueron lavadas una vez en medio de cultivo y se resuspendieron a 1×10^7 células/mL antes de ser usadas. Las células diana L540 a 1×10^6 células/mL en medio RPMI 1640 que contenía FBS al 10% (v/v) y probenecid 5 mM fueron marcadas con BATDA (bis(acetoximetil)2,2':6',2''-terpiridina-6,6''-dicarboxilato) 20 µM durante 20 minutos a 37°C. Las células diana fueron lavadas tres veces en PBS suplementado con HEPES 20 mM y probenecid 5 mM, resuspendidas a 1×10^5 células/mL y añadidas a las células efectoras en placas de 96 pocillos (1×10^4 células diana y 5×10^5 células efectoras/pocillo) con una relación final de diana a efector de 1:50. La liberación máxima se obtuvo mediante incubación de las células diana en Lysol al 3% (v/v) y la liberación espontánea obtenida solamente mediante incubación en medio de cultivo celular. Después de 1 h de incubación a 37°C, se recolectaron 20 µL de sobrenadante de cada pocillo y se añadieron a pocillos que contenían 180 µL de disolución de Europio. La reacción se leyó con un lector Perkin Elmer Fusion Alpha TRF usando un retardo de 400 µs y filtros de excitación y emisión de 330/80 y 620/10, respectivamente. Las cuentas por segundo se

representaron en función de la concentración de anticuerpo, y los datos fueron analizados mediante regresión no lineal, respuesta de dosis sigmoidal (de pendiente variable) usando el software GraphPad Prism 3.0. Se calculó el porcentaje de lisis específica del siguiente modo: (liberación experimental – liberación espontánea) / (liberación máxima – liberación espontánea) x 100. En todos los estudios, se incluyó el control de isotipo mAb humano y se comparó con los mAbs de MDX-060, LEX y LEX^{Opt}. Otros controles incluyeron células diana y efectoras sin mAb, células diana sin células efectoras y células diana y efectoras en presencia de Lysol al 3% (v/v).

Resultados

Expresión de mAb MDX-060 en el sistema LEX:

El MDX-060 es un anticuerpo anti-CD30 (anteriormente conocido como 5F11) que está siendo desarrollado para el tratamiento del linfoma de Hodgkins y el linfoma de célula grande anaplásica (Borchmann et al. (2003) Blood 102: 3737-3742). Se construyeron dos vectores binarios para la expresión de MDX-060 en el sistema LEX. El vector de expresión MDXA01 contenía genes de codón optimizado que codifican cadenas pesada (H) y ligera (L) de MDX-060, mientras que el vector MDXA04 contenía genes que codifican cadenas H y L además de un gen de ARNi quimérico de FucT/XylT (Figura 39). Se generaron líneas transgénicas independientes para ambos vectores de expresión, el MDXA01 (165 líneas) y el MDXA04 (195 líneas). Para mayor facilidad, en la discusión incluida a continuación los mAbs derivados de MDXA01 (*N*-glicosilación natural) y los mAbs derivados de MDXA04 (que contenían la construcción de ARNi de FucT/XylT) se denominarán MDX-060 LEX y MDX-060 LEX^{Opt}, respectivamente.

En primer lugar se escrutaron las plantas transgénicas para la expresión de mAb con un ELISA de IgG. Las líneas LEX^{Opt} con elevados niveles de expresión de mAb fueron ensayadas además para determinar la actividad de FucT y XylT. Las actividades de transferasa de la mayoría de las líneas MDX-060 LEX^{Opt} de elevada expresión fueron reducidas a los niveles del control negativo, lo que indica un silenciamiento efectivo en la mayoría de las líneas evaluadas (Figura 40). Las líneas MDX-060 LEX^{Opt} no exhibieron ninguna diferencia morfológica o de crecimiento en comparación con las plantas *Lemna* naturales (datos no mostrados).

Se determinaron las estabilidades térmicas de los mAbs MDX-060 CHO, LEX y LEX^{Opt} usando calorimetría de barrido diferencial (DSC). Los tres mAbs exhibieron curvas cinéticas de fusión similares (datos no mostrados) y temperaturas de punto de transición de fusión similares (Tabla 2, a continuación), demostrando aún más la integridad estructural de los mAbs MDX-060 LEX y LEX^{Opt} producidos en *Lemna* en comparación con el mAb MDX-060 CHO. El análisis de SDS-PAGE en condiciones no reductoras (Figura 41A) y reductoras (Figura 41B) demostró que los mAbs de las líneas MDX-060 LEX^{Opt} y MDX-060 CHO tenían perfiles de proteínas similares, siendo el mAb el componente principal del extracto proteínico.

Tabla 2. Comparación de las estabilidades térmicas de mAbs MDX-060 CHO, MDX-060 LEX y glico-optimizado MDX-060 LEX^{Opt} mediante calorimetría de barrido diferencia (DSC).

Anticuerpo	T _{m1} (°C)	T _{m2} (°C)	T _{m3} (°C)
MDX-060 CHO	72	75	84
MDX-060 LEX	71	75	84
MDX-060 LEX ^{Opt}	72	76	84

Estructuras de *N*-glicano de mAbs MDX-060 CHO, LEX y LEX^{Opt}.

Los perfiles de *N*-glicano de mAbs recombinantes derivados de MDX-060 CHO, MDX-060 LEX y MDX-060 LEX^{Opt} se determinaron mediante MALDI-TOF MS en modo reflectrón negativo y HPLC-QTOF MS de fase normal (NP). Las estructuras de los *N*-glicanos referidas en la siguiente discusión se muestran en la Figura 53.

El análisis MALDI-TOF MS de *N*-glicanos procedentes de las líneas MDX-060 CHO indicó la presencia de cuatro *N*-glicanos principales con valores de *m/z* correspondientes a GnGnF⁶ (nomenclatura derivada de <http://www.proglycan.com>), Man5, GnA_{iso}F⁶ y AAF⁶ marcados con 2-AA (Figura 42). La NP-HPLC separó el *N*-glicano GnA_{iso}F⁶ en sus dos isoformas (Gal unida al brazo α-1,6-Man o al brazo α-1,3-Man) elevando el número total de *N*-glicanos principales encontrados en MDX-060 CHO a cinco (Figura 43). No se llevó a cabo una fragmentación MS/MS de los picos para confirmar la identidad de cada isoforma; sin embargo, la mayor abundancia del primer pico sugirió que la Gal estaba unida en el brazo α-1,6-Man de dicho *N*-glicano (Shinkawa et al. (2003) J. Biol. Chem. 278: 3466-3473; Zhu et al. (2005) Nat. Biotechnol. 23: 1159-1169). El análisis QTOF MS en modo negativo en línea proporcionó valores *m/z* correspondientes a GnGnF⁶, Man5, GnA_{iso}F⁶ (ambas isoformas) y AAF⁶ doblemente cargados que confirman los resultados de MALDI-TOF MS (Tabla 3, a continuación). La integración de los picos de la traza fluorescente reveló que GnGnF⁶, Man5, AGnF⁶, GnAF⁶ y AAF⁶ constituían el 50,8, 2,5, 26,1, 10,7 y 6,8%, respectivamente, del total de estructuras de *N*-glicano de MDX-060 CHO. El 3,1% restante de *N*-glicanos era una

mezcla de GnGn, GnM_{iso}F⁶GnM_{iso} y MM, sin estar ninguna estructura por encima del 1,2% del total (datos no mostrados).

Tabla 3. Resumen de las masas observadas en MALDI-TOF y QTOF MS de los principales *N*-glicanos marcados con 2-AA procedentes de mAbs MDXA-060 producidos por líneas celulares de CHO (CHO), *Lemna* natural (LEX) o *Lemna* glico-optimizada que expresa la construcción de ARNi (LEX^{Opt}).

Nombre de <i>N</i> -glicano ^a	Estructura Propuesta ^b	m/z teórico		Observado en MALDI-TOF ^c	Observado en Q-TOF ^c	% de Área de Pico ^c
CHO		[M-H] ⁻	[M-2H] ²⁻	[M-H] ⁻	[M-2H] ²⁻	
GnGnF ⁶ -2AA		1582,90	790,7911	1582,455	790,7825	50,8
Man5-2AA		1354,479	676,7436	1354,392	676,7343	2,50
GnA _{iso} F ⁶ -2AA		1744,642	871,8175	1744,492	871,7970	36,8
AAF ⁶ -2AA		1906,695	952,8438	1906,567	952,8181	6,80
LEX						
GnGn-2AA		1436,532	717,7622	1436,549	717,7894	8,40
CHO						
GnGnX-2AA		1568,574	783,7833	1568,581	783,8150	17,2
GnGnXF ³ -2AA		1714,632	856,8122	1714,615	856,853	67,4
LEX^{Opt}						
GnGn-2AA		1436,532	717,7622	1436,523	717,7993	95,8

^aLos nombres de los *N*-glicanos se basan en la nomenclatura de Proglycan (<http://www.proglycan.com>). 2AA, ácido 2-aminobenzoico.

^bLos símbolos de las estructuras de *N*-glicano propuestas son los siguientes: , *N*-acetilglucosamina; , manosa; , galactosa; , xilosa; , α-1,3-fucosa; , α-1,6-fucosa; , ácido 2-aminobenzoico.

^cLos valores m/z y el % de área de pico de cada estructura de *N*-glicano se obtienen de la Figura 2.

El análisis MALDI-TOF MS del mAb MDX-060 LEX natural reveló la presencia de tres especies principales con valores m/z correspondientes a GnGnXF³, GnGnX y GnGn (Figura 42). El análisis NP-HPLC-QTOF MS mostro tres picos fluorescentes principales con valores m/z correspondientes a GnGnXF³, GnGnX y GnGn cargados doblemente, confirmando de nuevo los resultados de MALDI-TOF MS (Figura 43; Tabla 3). La integración de los picos fluorescentes indicó que GnGnXF³, GnGnX y GnGn constituían el 67,4, 17,2 y 8,4%, respectivamente, de los *N*-glicanos totales derivados del mAb MDX-060 LEX. Se determinó que el 7% restante de *N*-glicanos era una mezcla de MM, GnM_{iso}, MMXF³, GnGnF³, GnM_{iso}XF³, Man6, Man7, Gn(FA)_{iso}XF³, Man8 y Man9 sin ningún *N*-glicano por encima del 2% del total. Se observaron resultados similares para los mAbs aislados de dos líneas MDX-060

transformadas independientemente (datos no mostrados). El grupo sencillo de *N*-glicanos de los mAbs LEX demostrado en la presente memoria proporciona un punto de partida susceptible para la glico-optimización.

Al contrario que los mAb MDX-060 LEX, los *N*-glicanos procedentes de MDX-060 LEX^{Opt} poseían GnGn como la principal especie de *N*-glicano en ambos análisis, MALDI-TOF MS y NP-HPLC-QTOF MS (Figuras 42 y 43; Tabla 3). El GnGn constituyó el 95,8% del total de *N*-glicanos, siendo el 4,2% restante de los *N*-glicanos MM, GnM_{ISO}, GnA_{ISO}, Man6, Man7 y Man8 sin ninguna estructura que superara el 1,2% del total de *N*-glicanos. Ninguno de los *N*-glicanos de LEX^{Opt} contenía fucosa (Fuc) o xilosa (Xyl). Estos resultados demuestran que la co-expresión de construcción de ARNi dirigida a FucT y XylT de Lemna dio como resultado la eliminación completa de *N*-glicanos que contienen Fuc y Xyl de los mAbs MDX-060 LEX^{Opt} y produjo glicoformas de mAb altamente homogéneas. Los mismos resultados se obtuvieron para mAb MDX-060 LEX^{Opt} recolectado de una línea transgénica independiente (línea 225) que comprende el vector de expresión MDXA04, a una escala de crecimiento diferente (300 g de tejido para la línea transgénica 225 frente a 1 g de tejido para la línea transgénica 52, que produjo el mAb MDX-060 LEX^{Opt} que tiene el perfil de *N*-glicano mostrado en las Figuras 42 y 43. Al contrario que los sistemas de cultivo de células de mamífero en donde la heterogeneidad de *N*-glicanos puede cambiar con las condiciones de cultivo, la escala de crecimiento y el periodo de crecimiento (Kanda et al. (2006) Biotechnol. Bioeng. 94: 680-688), se demostró que la uniformidad de glicano observada con mAbs LEX^{Opt} era consistente en una serie de condiciones y escalas de crecimiento.

La ausencia de Fuc o Xyl en *N*-glicanos de mAb MDX-060 LEX^{Opt} se volvió a confirmar mediante análisis de monosacáridos (Tabla 4, a continuación). Los monosacáridos de mAbs MDX-060 CHO, LEX y LEX^{Opt} se liberaron mediante hidrólisis ácida y se analizaron mediante cromatografía de intercambio aniónico de alta resolución (HPAEC) acoplado a detección amperométrica pulsada (PAD). Las relaciones de monosacáridos para los residuos de Man y GlcNAc fueron similares para mAbs CHO y LEX naturales y se correlacionaron bien con los valores esperados. Los mAbs LEX presentaron valores significativamente reducidos de contenido de Gal y Fuc y presentaron un aumento de Xyl en comparación con los mAbs derivados de CHO. El análisis de monosacáridos de mAbs derivados de Lemna reveló que aunque había Fuc y Xyl presentes en los *N*-glicanos de LEX natural, no se detectaron en mAbs LEX^{Opt}. En conjunto, estos resultados demuestran que la co-expresión de una construcción de ARNi dirigida a XylT y FucT de Lemna da como resultado la eliminación de *N*-glicanos que contengan Fuc y Xyl a partir de mAbs MDX-060 LEX^{Opt}, y produce glicoformas de mAb altamente homogéneas. La robustez de esta estrategia de glico-optimización ha sido confirmada con múltiples líneas vegetales independientes que expresan el mAb MDX-060 LEX^{Opt}, así como con otros mAbs expresados en el sistema LEX, por ejemplo, el anticuerpo monoclonal mAbI discutido en los ejemplos mostrados anteriormente en la presente memoria. En las siguientes transformaciones, se han obtenido niveles de expresión de mAb glico-optimizado de hasta el 6% de la proteína soluble total (TSP).

Tabla 4. Monosacáridos liberados de mAbs MDX-060 CHO, LEX y LEX^{Opt} pro hidrólisis ácida y analizados mediante HPAEC-PAD. El contenido en monosacáridos de cada mAb se determinó normalizando respecto a los controles de carbohidratos.

Monosacárido	MDX-060 CHO pmol (% del total)	MDX-060 LEX pmol (% del total)	MDX-060 LEX ^{Opt} pmol (% del total)
Fuc	254 (20)	232 (13)	0
GlcNAc	605 (47)	773 (45)	1.003 (67)
Gal	75 (6)	0	0
Man	355 (27)	491 (29)	501 (33)
Xyl	0	226 (13)	0
Total	1.289 (100)	1.722 (100)	1.504 (100)

Actividad funcional de mAbs MDX-060 CHO, LEX y LEX^{Opt}.

Las propiedades de unión a antígeno de los mAbs MDX-060 CHO, MDX-060 LEX y MDX-060 LEX^{Opt} se determinaron usando células L540 que expresan CD30. Los tres mAbs presentaron curvas de unión casi idénticas (Figura 43). Se determinó que las concentraciones EC₅₀ eran 0,180 µg/mL, 0,227 µg/mL y 0,196 µg/mL para MDX-060 CHO, LEX y LEX^{Opt}, respectivamente (Figura 44), lo que indica que la unión a antígeno de los tres mAbs era similar.

Se ha demostrado que la función celular efectora mediada por receptor Fc es importante para la actividad *in vivo* de muchos mAbs terapéuticos. Puesto que el FcR expresado en células NK y macrófagos responsable de la actividad ADCC es FcRγIIa, se comparó la unión de varios mAbs a dicho receptor. La unión de mAbs CHO, LEX y LEX^{Opt} se

determinó a través de la unión de equilibrio de los mAbs con células efectoras que expresan dos alotipos de FcRyIIa humano diferentes (Phe¹⁵⁸ o Val¹⁵⁸). El MDX-060 LEX presentó un aumento de 1,7 veces en la unión de FcRyIIaPhe¹⁵⁸ y un descenso de 0,4 veces en la de FcRyIIaVal¹⁵⁸ en comparación con el mAb derivado de CHO, lo que demuestra que la unión a receptor correspondiente a los mAbs CHO y LEX era similar. Por el contrario, los mAbs LEX^{Opt} presentaron un aumento de 27 veces y 15 veces en la afinidad a FcRyIIaPhe¹⁵⁸ y FcRyIIaVal¹⁵⁸ respectivamente, en comparación con los mAbs CHO (Figura 45). Estos resultados sugieren que el silenciamiento de ARNi de las actividades de FucT y XylT de *Lemna* en las líneas LEX^{Opt} produjo mAbs con una unión a FcR potenciada.

Se determinaron las actividades de ADCC de los mAbs CHO, LEX y LEX^{Opt} incubando mAbs con células efectoras humanas homocigotas (FcRyIIaPhe¹⁵⁸) o heterocigotas (FcRyIIaPhe/Val¹⁵⁸) y células diana L540 marcadas con BATDA (bis(acetoximetil) 2,2':6',2"-terpiridina-6,6"-dicarboxilato) (Figura 45). Los mAbs MDX-060 LEX (31%) presentaron el mismo porcentaje de lisis máxima que los mAbs CHO (31%) usando células efectoras heterocigotas FcRyIIaPhe/Val¹⁵⁸ (Figura 46) con valores EC₅₀ similares. El porcentaje de lisis celular máximo para mAbs LEX (27%) aumentó ligeramente en comparación con los mAbs CHO (15%) usando células efectoras heterocigotas FcRyIIaPhe/Phe¹⁵⁸. Cabe destacar que los mAbs LEX^{Opt} presentaron una actividad ADCC significativamente incrementada en comparación con los mAbs MDX-060 CHO y LEX, independientemente del genotipo donante. Esto se determinó a través de un aumento de potencia y de eficacia. El porcentaje de lisis máxima para los MDX-060 LEX^{Opt} fue de 45% para ambos experimentos, mientras que el valor EC₅₀ fue entre 3 y 5 veces más bajo que para los mAbs MDX-060 LEX y MDX-060 CHO, respectivamente, para células efectoras FcRyIIa Val/Phe¹⁵⁸ y entre 2 y 3 veces más bajo para células efectoras FcRyIIa Phe/Phe¹⁵⁸. Estos resultados demuestran que la eliminación de *N*-glicanos que contienen Fuc y Xyl procedentes de mAbs MDX-060 LEX^{Opt} produjo un aumento de la actividad de ADCC y por tanto mejora su potencial terapéutico.

Análisis RP-HPLC-Q-TOF MS de IgG intacta para MDX-060 LEX y MDX-060 LEX^{Opt}.

Los IgGs purificados de Proteína A (50 µg) fueron desalinizados usando el sistema de HPCL Waters 2695 equipado con una columna Poros R1-10 (2 mm x 30 mm; Applied Biosystems). Los IgGs fueron detectados y analizados usando un detector UV (280 nm) de longitud de onda dual y el sistema Waters Q-TOF API US. Las separaciones se llevaron a cabo a 0,15 mL/min, 60°C, usando ácido trifluoroacético al 0,05% (v/v) (TFA; disolvente A) y TFA al 0,05% (v/v) y acetonitrilo al 80% (v/v) (disolvente B). La elución de muestras se llevó a cabo usando un aumento lineal desde 30 a 50% de B durante 5 minutos, un aumento hasta 80% de B durante 5 minutos. La relación de disolventes se mantuvo en 80% de B durante otros 4 minutos adicionales, seguido de un lavado con un 100% de B durante 1 minuto y un equilibrado de la columna con un 30% de B durante 15 minutos antes del siguiente experimento.

El análisis Q-TOF se llevó a cabo en modo de ión positivo con temperaturas de fuente y desolvatación de 100°C y 300°C, respectivamente, y voltajes de capilar y de cono de 3,0 y 60 V, respectivamente. Los datos son el resultado de combinar ≥ 100 escaneos individuales y de la deconvolución respecto al espectro de masas originales usando MaxEnt 1.

Véase también Triguero et al. (2005) Plant Biotechnol. J. 3: 449-457; Takahashi et al. (1998) Anal. Biochem. 255: 183-187; Dillon et al. (2004) J. Chromatogr. A. 1053: 299-305.

La Figura 47 muestra el análisis de masas intactas de las composiciones de mAb MDX-060 LEX producidas en *L. minor* natural que comprende la construcción MDXA01. Cuando las expresiones de XylT y FucT no son suprimidas en *L. minor*, la composición de mAb MDX-060 LEX producida recombinantemente comprende al menos 7 glicoformas diferentes, siendo la glicoforma G0XF³ la especie predominante. Nótese la ausencia de un pico que representa la glicoforma G0.

La Figura 48 muestra el análisis de masas de glicano de la cadena pesada del mAb MDX-060 LEX producido en *L. minor* natural que comprende la construcción MDXA01. Cuando las expresiones de XylT y FucT no son suprimidas en *L. minor*, las especies de *N*-glicano predominantes presentes en G0XF³, con picos principales adicionales que reflejan las especies G0X. Nótese la menor presencia de especies de glicano G0.

La Figura 49 muestra el análisis de masas intactas de las composiciones de mAb MDX-060 LEX^{Opt} producidas en *L. minor* transgénica que comprende la construcción MDXA04. Cuando las expresiones de XylT y FucT son suprimidas en *L. minor*, la composición de mAb intactos contiene solo *N*-glicanos G0. Adicionalmente, la composición es sustancialmente homogénea respecto a la glicoforma G0 (pico 2), en donde ambos sitios de glicosilación están ocupados por las especies de *N*-glicano G0, con dos picos menores que reflejan cantidades traza de glicoformas precursoras (pico 1, que muestra mAb que tiene una región Fc en donde el dominio C_H2 de una cadena pesada tiene una especie de glicano G0 unido a Asn 297, y el dominio C_H2 de la otra cadena pesada está sin glicosilar; y el pico 3, que muestra mAb que tiene una región Fc en la que el sitio de glicosilación Asn 297 de cada uno de los dominios C_H2 tiene una especie de glicano G0 unida, con una tercera especie de glicano G0 unida a un sitio de glicosilación adicional dentro de la estructura de mAb).

La Figura 50 muestra el análisis de masas de glicano de la cadena pesada del mAb MDX-060 LEX^{Opt} en *L. minor* transgénica que comprende la construcción MDXA04. Cuando las expresiones de XylT y FucT se suprimen en *L.*

minor, la única especie de *N*-glicano fácilmente detectable unida a los sitios de glicosilación Asn 297 de los dominios C_H2 de las cadenas pesadas es G0.

Discusión

La glico-optimización de MDX-060 se llevó a cabo mediante la co-expresión con una casete de ARNi dirigida a silenciar los genes endógenos de FucT y XylT en *Lemna*. Este silenciamiento simultáneo de ambos genes FucT y XylT se logró usando un único transcrito de ARNi. La ausencia de Fuc y Xyl sobre el mAb LEX^{Opt} se confirmó mediante análisis MALDI-TOF, NP-HPLC-QTOF MS y de monosacáridos de los *N*-glicanos purificados a partir del mAb MDX-060 LEX^{Opt}. Estos análisis corroboran la carencia de actividad de transferasa observada en membranas microsomales. De manera importante, >95% de los *N*-glicanos liberados de los mAbs LEX^{Opt} fueron de una única estructura, GnGn, lo que indica que esta estrategia tenía el beneficio añadido de producir mAbs con un perfil de *N*-glicanos homogéneo. Se observó que los mAbs MDX-060 LEX y LEX^{Opt} eran indistinguibles en lo que respecta a la estabilidad térmica y la unión a antígeno, en comparación con MDX-060 CHO. También se observó que el análisis electroforético era similar para los tres mAbs. De hecho, las únicas diferencias estructurales detectadas fueron en los perfiles de *N*-glicano de mAb.

Sin pretender establecer una teoría, la capacidad de las líneas de MDX-060 LEX^{Opt} para producir mAbs con una única especie de *N*-glicano principal puede basarse en la distribución de glicoformas de mAb más uniforme observada en la *Lemna* natural. Los *N*-glicanos liberados de mAbs purificados a partir de tabaco (Fujiyama et al. (2006) J. Biosci. Bioeng. 101: 212-218; Elbers et al. (2001) Plant Physiol. 126: 1314-1322; Bakker et al. (2001) Proc. Natl. Acad. Sci. 98: 2899-2904), alfalfa (Bardor et al. (2003) Plant Biotechnol. J. 1: 451-462), y musgo (Koprivova et al. (2003) Plant Bio. 5: 582-591) naturales demuestran que la heterogeneidad de glicoforma de mAb en plantas con *N*-glicosilación natural puede oscilar entre cinco estructura principales diferentes (alfalfa) (Bardor et al. (2003) Plant Biotechnol. J. 1: 451-462) hasta ocho (tabaco) (Fujiyama et al. (2006) J. Biosci. Bioeng. 101: 212-218). El MDX-060 LEX posee sólo tres estructuras de *N*-glicano principales (GnGn, GnGnX y GnGnXF³). Esta disposición sencilla de *N*-glicanos en mAbs producida por la *Lemna* natural puede proporcionar un punto de partida más susceptible para la glico-optimización que conduce a una mayor homogeneidad que la observada en otros sistemas.

Se ha demostrado que la función de célula efectora mediada por receptor Fc es importante para la actividad *in vivo* de muchos mAbs terapéuticos. En este estudio se comparó la actividad ADCC de mAbs MDX-060 CHO, MDX-060 LEX y MDX-060 LEX^{Opt}. Puesto que el FcR expresado en células NK y macrófagos responsables de la actividad ADCC es el FcγRIIIa, también se comparó la unión de los diversos mAbs a dicho receptor. Los resultados discutidos anteriormente demuestran que el mAb MDX-060 LEX^{Opt} tenía una afinidad de unión incrementada (de 15 a 25 veces) y una unión máxima (4-5 veces) a FcγRIIIa, así como una actividad ADCC mejorada en comparación con los mAbs MDX-060 CHO y MDX-060 LEX. Se ha demostrado previamente que la eliminación de Fuc α-1,6-ligada de varios mAbs producidos en otros sistemas de expresión aumenta la unión a FcR y potencia la función de ADCC (Shinkawa et al. (2003) J. Biol. Chem. 278: 3466-3473; Shields et al. (2002) J. Biol. Chem. 277: 26733-26740; Niwa et al. (2004) Clin. Cancer Res. 10: 6248-6255). Los resultados presentados en la presente memoria sugieren que la eliminación de la Fuc α-1,3-ligada procedente de mAbs MDX-060 LEX^{Opt} tiene el mismo efecto sobre la función del mAb que la eliminación de Fuc α-1,6-ligada.

En este estudio se evaluaron dos isoformas polimórficas naturales del FcγRIIIa en el residuo 15841, Val¹⁵⁸ y Phe¹⁵⁸. El MDX-060 LEX^{Opt} muestra una mayor afinidad de unión a FcγRIIIa-Val¹⁵⁸ en comparación con el FcγRIIIa-Phe¹⁵⁸, como se ha observado con otros mAbs de IgG1 (Shields et al. (2002) J. Biol. Chem. 277: 26733-26740). El hecho de que se observe un aumento de la unión con MDX-060 LEX^{Opt} con ambas isoformas es importante ya que se observó que la unión diferencial a Val¹⁵⁸ sobre Phe¹⁵⁸ era predictiva de capacidad de respuesta clínica e inmunológica en determinados grupos de pacientes que reciben tratamiento anti-CD20 (Cartron et al. (2002) Blood 99: 754-758; Weng et al. (2003) J. Clin. Oncol. 21: 3940-3947; Weng et al. (2004) J. Clin. Oncol. 22: 4717-4724). Se ha planteado la hipótesis de que este incremento de la unión da como resultado un mayor porcentaje de pacientes que responden al tratamiento que requiere funcionalidad Fc.

También se observó un incremento similar de la actividad de ADCC. En este estudio, el mAb MDX-060 LEX^{Opt} demostró un aumento de la lisis celular y un descenso del valor de EC₅₀, dando como resultado un aumento de la eficacia y la potencia cuando se compara con MDX-060 CHO. Esto corresponde a un aumento de 20 veces en la actividad, determinado tomando el porcentaje máximo de lisis de MDX-060 CHO y calculando la concentración de mAb MDX-060 LEX^{Opt} que da lugar al mismo porcentaje de lisis celular. Como en el estudio de unión a FcγRIIIa, el aumento de actividad de ADCC se observó en un donante de célula efectora homocigota FcγRIIIa-Phe/Phe¹⁵⁸ y una heterocigota FcγRIIIa-Phe/Val¹⁵⁸. Los resultados presentados aquí sugieren que la eliminación del Fuc α-1,3-ligada de los mAbs MDX-060 LEX^{Opt} tiene el mismo efecto sobre la función de mAb que la eliminación de Fuc α-1,6-ligada.

La robustez de esta estrategia de glico-optimización se ha demostrado con múltiples líneas vegetales de *Lemna* independientes que expresan el mAb MDX-060 LEX^{Opt}, así como con otros mAbs expresados en el sistema de expresión de *Lemna* (véanse, por ejemplo, los Ejemplos 2-4 anteriores). Además, no existe una diferencia aparente en el fenotipo de planta o en la tasa de crecimiento en comparación con plantas de *Lemna* naturales. Al contrario que los sistemas de cultivo de células de mamífero en los que la heterogeneidad de *N*-glicanos puede cambiar con las condiciones de cultivo, la escala de crecimiento y el periodo de crecimiento⁸, se ha demostrado que la

uniformidad de glicano observada con mAbs LEX^{Opt} es consistente bajo una variedad de condiciones y escalas de crecimiento (datos no mostrados).

En conclusión, se ha usado una estrategia de ARNi para producir anticuerpos anti-CD30 glico-optimizados en el sistema de expresión *Lemna*. El mAb resultante consta de una estructura de *N*-glicano principal individual, sin ninguna evidencia de los residuos Fuc y Xyl específicos de planta. Además, el mAb optimizado resultante tiene una actividad de ADCC y una actividad de unión a FcγR1IIa incrementadas en comparación con el mAb derivado de CHO. El perfil de glicosilación homogénea obtenido en los mAbs producidos en un sistema de expresión de *Lemna* que tiene la estrategia de bloqueo de gen FucT + XylT hace posible expresar dichos mAbs con una consistencia de producción mejorada.

Ejemplo 7: Aumento de Escala de Producción de mAb1 de Glicano Optimizado en *Lemna minor*

La línea 24 de *L. minor* transgénica que comprende la construcción mAb104 de la Figura 12 (que proporciona la supresión de FucT y XylT) fue generada de un modo similar al descrito anteriormente. Después de su generación, la línea transgénica 24 se mantuvo continuamente mediante cultivo clonal, en donde una submuestra del cultivo vegetal era transferida periódicamente a medio de cultivo fresco para continuar con el cultivo. Esta línea transgénica fue analizada para determinar el modelo de *N*-glicosilación del anticuerpo mAb1 producido recombinantemente después de un aumento en la escala de producción desde 1 g de tejido hasta 300 g (0,3 kg) de tejido, y un aumento adicional de la escala de producción hasta 6,5 kg de tejido. El proceso de escalado de la producción hasta 6,5 kg de tejido se produjo aproximadamente 8 meses después de generar la línea transgénica 24. Los resultados se muestran en las Figuras 51A (análisis MALDI-TOF de *N*-glicanos) y 51B (análisis de fluorescencia de HPLC de *N*-glicanos).

El perfil de glicosilación correspondiente al anticuerpo mAb1 producido mediante la línea transgénica 24 que comprende la construcción mAb104 permaneció homogéneo en el aumento de escala de producción de 1 g de tejido a 0,3 kg de tejido, y en el aumento adicional de escala de producción hasta 6,5 kg de tejido, y por tanto se caracterizó por la presencia de un único pico predominante correspondiente a la especie de glicano GnGn (es decir, G0). Por tanto, la homogeneidad del perfil de glicosilación en *L. minor* transgénica que comprende la construcción mAb104 se mantuvo consistentemente con un incremento de aproximadamente 6.500 veces en la escala de producción (es decir, desde 1 g hasta 6,5 kg). Además, la homogeneidad del perfil de glicosilación se mantuvo consistentemente en dicha línea transgénica a los 8 meses después de su generación.

Estos datos demuestran que la homogeneidad del perfil de glicosilación en *L. minor* transgénica que comprende la construcción mAb104 permanece consistente con un aumento de la escala de producción de al menos 6.500 veces. Además, la homogeneidad del perfil de glicosilación en *L. minor* transgénica que contiene la construcción mAb104 se mantiene durante al menos 8 meses después de la generación de la línea transgénica. Cabría esperar que la homogeneidad del perfil de glicosilación se mantuviera con un aumento adicional de la escala de producción, y por tanto, por ejemplo, cabría esperar que se mantuviera si la escala de producción se incrementara otras 4 veces por encima de 6,5 kg (por ejemplo, un escalado desde 6,5 kg hasta 26 kg). Cabría esperar que la homogeneidad del perfil de glicosilación se mantuviera con un cultivo clonal continuo de la línea transgénica bastante más allá de 8 meses después de la generación de la línea transgénica.

Ejemplo 8: Modelo de Glicosilación correspondiente a Glicoproteínas Endógenas en Líneas de *Lemna minor* Transgénica para la Construcción de ARNi mAb104

La proteasa L40 es una glicoproteína endógena representativa producida en *L. minor*. Con el objetivo de determinar el impacto de la construcción de ARNi mAb104 (Figura 12) sobre la glicosilación de proteínas endógenas, se aisló la proteína L40 a partir de una línea transgénica de *L. minor* para la construcción de ARNi mAb104 usando cromatografía de afinidad de benzamidina. Se analizó el modelo de *N*-glicosilación correspondiente a la proteína L40 aislada usando análisis MALDI-TOF de la manera descrita anteriormente. Los resultados se muestran en la Figura 52.

Como se puede ver en este análisis, la supresión de las expresiones de FucT y XylT usando la construcción de ARNi mAb104 da como resultado glicoproteínas endógenas que tienen un modelo de glicosilación homogéneo consistente con el observado para glicoproteínas recombinantes. Por tanto, el perfil de *N*-glicanos heterogéneo correspondiente a la glicoproteína L40 aislada de *L. minor* que tiene la maquinaria de glicosilación natural está representado por una mezcla de especies de *N*-glicano que tienen el residuo de xilosa β1,2-ligada, o ambos residuos, de xilosa β1,2-ligada y de fucosa de núcleo α1,3-ligada, ligados. Por el contrario, el perfil de *N*-glicanos homogéneo correspondiente a L40 aislada de *L. minor* transgénica para la construcción de ARNi mAb104 está representada por un único pico predominante correspondiente a especies de glicano G0, y se caracteriza por la ausencia de especies de *N*-glicano que tienen el residuo de xilosa β1,2-ligada, o ambos residuos, de xilosa β1,2-ligada y de fucosa de núcleo α1,3-ligada, ligados.

Ejemplo 9: Producción de Anticuerpos Monoclonales Anti-CD20 y Anti-HER2 que tienen una Actividad de ADCC Incrementada

IDEC-C2B8 (IDEC Pharmaceuticals Corp., San Diego, California; disponible comercialmente bajo el nombre comercial Rituxan®, también denominado rituximab; véase la Patente de EE.UU. N° 5.736.137, incorporada a la

presente memoria a modo de referencia) es un anticuerpo monoclonal anti-CD20 quimérico que contiene regiones constantes de IgG1 y kappa humanas con regiones variables murinas aisladas a partir de un anticuerpo monoclonal anti-CD20 murino, IDEC-2B8 (Reff et al. (1984) Blood 83: 435-445). El rituximab tiene licencia para el tratamiento de linfoma de grado bajo de célula B o linfoma folicular no de Hodgkin (NHL). El anticuerpo anti-CD20 comercializado como rituximab (Rituxan®) se produce recombinantemente en células CHO. El modelo de glicosilación de este anticuerpo anti-CD20 expresado en CHO revela una mezcla heterogénea de glicoformas.

Un anticuerpo anti-ERBB2 humanizado se encuentra disponible comercialmente bajo el nombre comercial de Herceptin® (Genentech, Inc., San Francisco, California) (véase la Patente de EE.UU. N° 6.165.464). Dicho anticuerpo monoclonal humanizado recombinante tiene una elevada afinidad para p185HER2. Los ensayos clínicos iniciales con pacientes que presentan carcinomas de pecho metastáticos extensivos demuestran la capacidad de este anticuerpo monoclonal para inhibir el crecimiento de células de cáncer de mama que sobreexpresan HER2 (Baselga et al. (1996) J. Clin. Oncol. 14(3): 737-744).

Un anticuerpo de secuencia de rituximab y un anticuerpo de secuencia anti-ERBB2 de Herceptin® son producidos recombinantemente en *Lemna* que tiene un modelo de glicosilación natural, usando la construcción mAbI01 descrita anteriormente, y en *Lemna* que está modificada genéticamente para suprimir la expresión de FucT y XylIT, usando la construcción mAbI04 de ARNi quimérico, con las secuencias codificadoras de cadena pesada y ligera de rituximab o anti-ERRB2 reemplazando a las del mAb en cada una de dichas construcciones. Las secuencias que codifican las cadenas pesada y ligera opcionalmente son de codón optimizado con codones preferidos de *Lemna*. El mAb de secuencia de rituximab (es decir, que tiene la secuencia de aminoácidos del anticuerpo rituximab) y el mAb de secuencia de anti-ERRB2 (es decir, que tiene la secuencia de aminoácidos del mAb anti-ERRB2 Herceptin®) secretados son analizados para determinar el modelo de glicosilación como se ha descrito anteriormente en el Ejemplo 1.

El perfil de glicosilación correspondiente a mAb de secuencia de rituximab o a mAb de secuencia anti-ERRB2 intactos producidos en *Lemna* natural que comprende la construcción de tipo mAbI01 muestra un perfil heterogéneo con numerosos picos correspondientes a múltiples glicoformas. Por el contrario, el perfil de glicosilación correspondiente a mAb de secuencia de rituximab o a mAb de secuencia anti-ERRB2 intactos producidos en *Lemna* transgénica que comprende la construcción de tipo mAbI04 muestra una composición de glicoproteínas sustancialmente homogénea, con tres picos de glicoformas principales, el mayor de los cuales corresponde a la glicoforma G0, y los dos picos menores corresponden a cantidades traza de glicoformas precursoras, en donde los residuos de xilosa y fucosa no están unidos.

Las composiciones de anticuerpos monoclonales de secuencia de rituximab y de secuencia anti-ERBB2 que tienen un perfil de glicosilación que es sustancialmente homogéneo para la glicoforma G0 son evaluadas para determinar la actividad de ADCC y la unión a receptor de Fcγ IIIa (FcγRIIIa) en células NK humanas recién aisladas. Los mAbs de secuencia de rituximab y de secuencia anti-ERBB2 producidos en líneas de *L. minor* modificadas con la construcción de tipo mAbI04 exhiben la unión mejorada esperada en vista de la falta de residuos de fucosa, respecto a la unión observada para los mAbs de secuencia de rituximab y de secuencia de anti-ERBB2 producidos en líneas de *L. minor* que tienen la maquinaria de glicosilación natural (es decir, sin silenciamiento de FucT o XylIT). Además, la afinidad de unión es al menos fuerte para el mAb correspondiente producido en líneas celulares de CHO.

Se determina la actividad de ADCC de la glicoforma G0 sustancialmente homogénea del mAb de secuencia de rituximab y del mAb de secuencia de anti-ERBB2 producidos en la línea de *L. minor* modificada con la construcción de tipo mAbI04 usando células mononucleares sanguíneas periféricas humanas purificadas como células efectoras (véase, por ejemplo, Shinkawa et al. (2003) J. Biol. Chem. 278 (5): 3466-3473). La actividad de la glicoforma G0 de las composiciones de mAb de secuencia de rituximab y de secuencia anti-ERBB2 se ve mejorada en 50-1000 veces sobre la exhibida por los respectivos mAb de secuencia de rituximab o de secuencia de anti-ERBB2 producidos en la línea *L. minor* que tiene la maquinaria de glicosilación natural o producidos en células CHO.

La actividad de CDC de la glicoforma G0 sustancialmente homogénea de mAb de secuencia de rituximab producido en la línea de *L. minor* modificada con la construcción de ARNi de tipo mAbI04 es determinada usando ensayos estándares conocidos en la técnica (véase, por ejemplo, los ensayos de activación de complemento descritos en la Publicación de Solicitud de Patente de EE.UU. N° 2004/0167319), y se comparó con la observada para rituximab. Anticuerpos no relevantes actúan como control negativo. En un ensayo de este tipo, se mide la actividad de CDC de los diversos anticuerpos frente a células Daudi determinando la permeabilidad de membrana elevada usando un ensayo de exclusión de yoduro de propidio (PI), con suero de voluntarios sanos como fuente de complemento. El suero para lisis de complemento se prepara llevando sangre de voluntarios sanos a tubos vacutainer de gel autosep y activador de coagulación (BD Biosciences, Rutherford, Nueva Jersey), que se mantuvieron a temperatura ambiente durante 30-60 minutos y a continuación se centrifugaron a 3.000 rpm durante 5 minutos. Se recolectó el suero y se almacenó a -80°C.

En resumen, para este ensayo de actividad CDC, las células Daudi fueron lavadas y resuspendidas en RPMI-BSA al 1% a 1×10^6 células/mL. A las células Daudi se añaden concentraciones variables de la glicoforma G0 sustancialmente homogénea de mAb de secuencia de rituximab, rituximab y mAb de control negativo, y se permite la unión durante 10-15 minutos a temperatura ambiente. Después de esto, se añade suero como fuente de

- complemento hasta una concentración final de 20% (v/v) y la mezclas se incuban durante 45 minutos a 37°C. A continuación las células se mantienen a 4°C hasta el análisis. Entonces se añade cada muestra (150 µL) a 10 µL de disolución de PI (10 µg/mL en PBS) en un tubo FACS. La mezcla se analiza inmediatamente para lisis celular (número de células positivas a PI) mediante un citómetro de flujo FACScalibur y se analizan usando el software CellQuest pro (BD Biosciences, Mountain View, California). Se recogen al menos 5.000 eventos para análisis excluyendo los restos celulares mediante ajuste del promedio de dispersión hacia los lados y hacia delante (FCS, del inglés "forward sideward scatter").
- 5 La actividad de CDC de la glicoforma G0 sustancialmente homogénea de mAb de secuencia de rituximab se ve disminuida respecto a la exhibida por rituximab.
- 10 Todas las publicaciones y solicitudes de patente mencionadas en la especificación son indicativas del nivel de los especialistas en la técnica a la cual pertenece esta invención.

LISTADO DE SECUENCIAS

<110> Dickey, Lynn F. Cox, Kevin M. Peele, Charles G.

<120> COMPOSICIONES Y MÉTODOS PARA LA HUMANIZACIÓN Y OPTIMIZACIÓN DE N-GLICANOS EN PLANTAS

5 <130> 040989/322364

<150> 60/759,298

<151> 2006-01-17

10 <150> 60/790,373

<151> 2006-04-07

<150> 60/791,178

<151> 2006-04-11

15

<150> 60/812,702

<151> 2006-06-09

<150> 60/836,998

20 <151> 2006-08-11

<150> 60/860,358

<151> 2006-11-21

25 <160> 29

<170> FastSEQ for Windows Version 4.0

<210> 1

30 <211> 1865

<212> ADN

<213> Lemna minor

<220>

35 <221> CDS

<222> (243)...(1715)

<400> 1

```

acgcgggggg aagtgggtga gtagctcagt ggaaaattgg aaatgtctat tagaggggga 60
agagggggagg gatccgaggg gaacgaggaa ggtgtgccga attctcgtag atttcttcaa 120
ttcctgcaga tctcgtcttc tctctgattt cttcccgagc tccgcccgta ggaactcaat 180
cggactcgat ccaagttgac gaggcctacy gaggaaggcg attttccgaa gccctgcgat 240
cg atg gcc acc tct gct gct ggt gct ctc aac gcc ggt ggc agg gtc 287
  Met Ala Thr Ser Ala Ala Gly Ala Leu Asn Ala Gly Gly Arg Val
    1           5           10          15

ggg ggc agg agg agt tgg gtc aga ttg ctt ccc ttc ttt gtg ttg atg 335
Gly Gly Arg Arg Ser Trp Val Arg Leu Leu Pro Phe Phe Val Leu Met
          20           25           30

ctg gtg gta ggg gag atc tgg ttc ctc ggg cgg ctg gat gtg gtc aag 383
Leu Val Val Gly Glu Ile Trp Phe Leu Gly Arg Leu Asp Val Val Lys
          35           40           45

aac gcc gct atg gtt caa aac tgg act tcc tcc cac ttg ttt ttc tta 431
Asn Ala Ala Met Val Gln Asn Trp Thr Ser Ser His Leu Phe Phe Leu
          50           55           60

cca gtt tct tcc tac acg tgg tcc gag acc gtc aag gag gaa gag gat 479
Pro Val Ser Ser Tyr Thr Trp Ser Glu Thr Val Lys Glu Glu Glu Asp

```


ES 2 395 842 T3

65	70	75	
tgc aag gac tgg ctg gaa aga gta gat gcg gtc gat tac aag aga gat Cys Lys Asp Trp Leu Glu Arg Val Asp Ala Val Asp Tyr Lys Arg Asp 80 85 90 95			527
ttc cgt gtg gaa ccc gtt ctg gta aat gac gct gaa cag gat tgg agt Phe Arg Val Glu Pro Val Leu Val Asn Asp Ala Glu Gln Asp Trp Ser 100 105 110			575
tca tgt tca gtg ggc tgt aag ttc gga tca ttc ccc gga aga acg cct Ser Cys Ser Val Gly Cys Lys Phe Gly Ser Phe Pro Gly Arg Thr Pro 115 120 125			623
gat gct aca ttt ggt ttc tct cag aat cca tca aca gtc agt gtc cat Asp Ala Thr Phe Gly Phe Ser Gln Asn Pro Ser Thr Val Ser Val His 130 135 140			671
cga tcc atg gaa tca tcc cat tat tat ttg gag aat aat ctt gat aat Arg Ser Met Glu Ser Ser His Tyr Tyr Leu Glu Asn Asn Leu Asp Asn 145 150 155			719
gca cga cgg aaa ggc tat caa att gtg atg aca act agt ctc ttg tca Ala Arg Arg Lys Gly Tyr Gln Ile Val Met Thr Thr Ser Leu Leu Ser 160 165 170 175			767
gat gtg cct gtc ggt tat ttc tca tgg gct gaa tat gat atc atg gcg Asp Val Pro Val Gly Tyr Phe Ser Trp Ala Glu Tyr Asp Ile Met Ala 180 185 190			815
cct ctt cag ccg aaa act gct ggt gca ctt gct gct gca ttt ata tct Pro Leu Gln Pro Lys Thr Ala Gly Ala Leu Ala Ala Phe Ile Ser 195 200 205			863
aat tgc gga gca cgt aat ttc cgc ttg cag gcc ctt gat atg ctc gaa Asn Cys Gly Ala Arg Asn Phe Arg Leu Gln Ala Leu Asp Met Leu Glu 210 215 220			911
aag tgc aat att aag att gat tca tat ggt gct tgc cat cgc aac caa Lys Ser Asn Ile Lys Ile Asp Ser Tyr Gly Ala Cys His Arg Asn Gln 225 230 235			959
gac ggt aaa gtg gac aag gta caa act ttg aag cgg tat aag ttc agc Asp Gly Lys Val Asp Lys Val Gln Thr Leu Lys Arg Tyr Lys Phe Ser 240 245 250 255			1007
tta gct ttt gaa aac tgc aac gag gat gac tat gtt act gag aag ttc Leu Ala Phe Glu Asn Ser Asn Glu Asp Asp Tyr Val Thr Glu Lys Phe 260 265 270			1055
ttt caa tct ctt gtc gct gga gct att cct gtt gtc gtc gga gcc ccc Phe Gln Ser Leu Val Ala Gly Ala Ile Pro Val Val Val Gly Ala Pro 275 280 285			1103
aac att caa aat ttt gcg cca tct tct gat tca att ctg cac atc agg Asn Ile Gln Asn Phe Ala Pro Ser Ser Asp Ser Ile Leu His Ile Arg 290 295 300			1151
gag ccc aag gat gtc agt tca gtc gct gag aga atg aaa ttt ctc gct Glu Pro Lys Asp Val Ser Ser Val Ala Glu Arg Met Lys Phe Leu Ala 305 310 315			1199
tca aat cca gaa gca tat aac caa tca ctg agg tgg aag ttt gag ggc Ser Asn Pro Glu Ala Tyr Asn Gln Ser Leu Arg Trp Lys Phe Glu Gly 320 325 330 335			1247

ES 2 395 842 T3

```

cct tct aac tcc ttc aaa gcc ctg gtg gac atg gca gca gtt cac tcc 1295
Pro Ser Asn Ser Phe Lys Ala Leu Val Asp Met Ala Ala Val His Ser
340 345 350

tcc tgc cgc cta tgc att cac att gcc acc aag atc aga gag aag gaa 1343
Ser Cys Arg Leu Cys Ile His Ile Ala Thr Lys Ile Arg Glu Lys Glu
355 360 365

gag aga aac ccg aat ttc aag act cgc cct tgc aag tgc acc cgc aat 1391
Glu Arg Asn Pro Asn Phe Lys Thr Arg Pro Cys Lys Cys Thr Arg Asn
370 375 380

ggg tct acc tta tat cac tta tac gcc cgc gaa aga ggc acc ttt gac 1439
Gly Ser Thr Leu Tyr His Leu Tyr Ala Arg Glu Arg Gly Thr Phe Asp
385 390 395

ttc tta tca atc ttc atg aga tcg gat aat cta tca ctg aaa gcg ctg 1487
Phe Leu Ser Ile Phe Met Arg Ser Asp Asn Leu Ser Leu Lys Ala Leu
400 405 410 415

ggg tca aca gtt ctt gag aaa ttc agt tct ttg aag cac gtg ccg att 1535
Gly Ser Thr Val Leu Glu Lys Phe Ser Ser Leu Lys His Val Pro Ile
420 425 430

tgg aag aag gag agg cca gag agt ctg aaa gga ggg agc aag ctg gat 1583
Trp Lys Lys Glu Arg Pro Glu Ser Leu Lys Gly Gly Ser Lys Leu Asp
435 440 445

ctt tac aga atc tat cca gtg ggc att act cag aga gaa gct ctc ttc 1631
Leu Tyr Arg Ile Tyr Pro Val Gly Ile Thr Gln Arg Glu Ala Leu Phe
450 455 460

tct ttc cag ttc aac act gac aaa gaa ctt caa atc tac ctt gaa tcc 1679
Ser Phe Gln Phe Asn Thr Asp Lys Glu Leu Gln Ile Tyr Leu Glu Ser
465 470 475

cat cca tgt gcg aag ttt gaa gtc atc ttt att tga tccctgaggt 1725
His Pro Cys Ala Lys Phe Glu Val Ile Phe Ile *
480 485 490

aattaggtca cgaattcagc taatttggtt aattatgctt caagcccaca tggattttca 1785
tatcattaat tgaaggcata gttagttgat attgacattt tcgtctagga tcattctaaa 1845
gtctatccca atgaacttaa 1865

```

<210> 2

<211> 1473

5 <212> ADN

<213> Lemna minor

<220>

<221> CDS

10 <222> (1)...(1473)

<223> Codifica alfa-1, 3-fucosiltransferasa

<400> 2

ES 2 395 842 T3

atg gcc acc tct gct gct ggt gct ctc aac gcc ggt ggc agg gtc ggg	48
Met Ala Thr Ser Ala Ala Gly Ala Leu Asn Ala Gly Gly Arg Val Gly	
1 5 10 15	
ggc agg agg agt tgg gtc aga ttg ctt ccc ttc ttt gtg ttg atg ctg	96
Gly Arg Arg Ser Trp Val Arg Leu Leu Pro Phe Phe Val Leu Met Leu	
20 25 30	
gtg gta ggg gag atc tgg ttc ctc ggg cgg ctg gat gtg gtc aag aac	144

ES 2 395 842 T3

Val	Val	Gly	Glu	Ile	Trp	Phe	Leu	Gly	Arg	Leu	Asp	Val	Val	Lys	Asn	
		35					40					45				
gcc	gct	atg	gtt	caa	aac	tgg	act	tcc	tcc	cac	ttg	ttt	ttc	tta	cca	192
Ala	Ala	Met	Val	Gln	Asn	Trp	Thr	Ser	Ser	His	Leu	Phe	Phe	Leu	Pro	
		50				55					60					
gtt	tct	tcc	tac	acg	tgg	tcc	gag	acc	gtc	aag	gag	gaa	gag	gat	tgc	240
Val	Ser	Ser	Tyr	Thr	Trp	Ser	Glu	Thr	Val	Lys	Glu	Glu	Glu	Asp	Cys	
		65			70					75					80	
aag	gac	tgg	ctg	gaa	aga	gta	gat	gcg	gtc	gat	tac	aag	aga	gat	ttc	288
Lys	Asp	Trp	Leu		Arg	Val	Asp	Ala	Val	Asp	Tyr	Lys	Arg	Asp	Phe	
				85					90						95	
cgt	gtg	gaa	ccc	gtt	ctg	gta	aat	gac	gct	gaa	cag	gat	tgg	agt	tca	336
Arg	Val	Glu	Pro	Val	Leu	Val	Asn	Asp	Ala	Glu	Gln	Asp	Trp	Ser	Ser	
			100					105					110			
tgt	tca	gtg	ggc	tgt	aag	ttc	gga	tca	ttc	ccc	gga	aga	acg	cct	gat	384
Cys	Ser	Val	Gly	Cys	Lys	Phe	Gly	Ser	Phe	Pro	Gly	Arg	Thr	Pro	Asp	
		115					120					125				
gct	aca	ttt	ggg	ttc	tct	cag	aat	cca	tca	aca	gtc	agt	gtc	cat	cga	432
Ala	Thr	Phe	Gly	Phe	Ser	Gln	Asn	Pro	Ser	Thr	Val	Ser	Val	His	Arg	
		130				135					140					
tcc	atg	gaa	tca	tcc	cat	tat	tat	ttg	gag	aat	aat	ctt	gat	aat	gca	480
Ser	Met	Glu	Ser	Ser	His	Tyr	Tyr	Leu	Glu	Asn	Asn	Leu	Asp	Asn	Ala	
		145			150				155						160	
cga	cgg	aaa	ggc	tat	caa	att	gtg	atg	aca	act	agt	ctc	ttg	tca	gat	528
Arg	Arg	Lys	Gly	Tyr	Gln	Ile	Val	Met	Thr	Thr	Ser	Leu	Leu	Ser	Asp	
				165				170						175		
gtg	cct	gtc	ggg	tat	ttc	tca	tgg	gct	gaa	tat	gat	atc	atg	gcg	cct	576
Val	Pro	Val	Gly	Tyr	Phe	Ser	Trp	Ala	Glu	Tyr	Asp	Ile	Met	Ala	Pro	
			180					185					190			
ctt	cag	ccg	aaa	act	gct	ggg	gca	ctt	gct	gct	gca	ttt	ata	tct	aat	624
Leu	Gln	Pro	Lys	Thr	Ala	Gly	Ala	Leu	Ala	Ala	Ala	Phe	Ile	Ser	Asn	
		195					200					205				
tgc	gga	gca	cgt	aat	ttc	cgc	ttg	cag	gcc	ctt	gat	atg	ctc	gaa	aag	672
Cys	Gly	Ala	Arg	Asn	Phe	Arg	Leu	Gln	Ala	Leu	Asp	Met	Leu	Glu	Lys	
		210				215					220					
tcg	aat	att	aag	att	gat	tca	tat	ggg	gct	tgc	cat	cgc	aac	caa	gac	720
Ser	Asn	Ile	Lys	Ile	Asp	Ser	Tyr	Gly	Ala	Cys	His	Arg	Asn	Gln	Asp	
					230					235					240	
ggg	aaa	gtg	gac	aag	gta	caa	act	ttg	aag	cgg	tat	aag	ttc	agc	tta	768
Gly	Lys	Val	Asp	Lys	Val	Gln	Thr	Leu	Lys	Arg	Tyr	Lys	Phe	Ser	Leu	
				245					250					255		
gct	ttt	gaa	aac	tcg	aac	gag	gat	gac	tat	gtt	act	gag	aag	ttc	ttt	816
Ala	Phe	Glu	Asn	Ser	Asn	Glu	Asp	Asp	Tyr	Val	Thr	Glu	Lys	Phe	Phe	
			260					265					270			
caa	tct	ctt	gtc	gct	gga	gct	att	cct	gtt	gtc	gtc	gga	gcc	ccc	aac	864
Gln	Ser	Leu	Val	Ala	Gly	Ala	Ile	Pro	Val	Val	Val	Gly	Ala	Pro	Asn	
			275				280						285			
att	caa	aat	ttt	gcg	cca	tct	tct	gat	tca	att	ctg	cac	atc	agg	gag	912
Ile	Gln	Asn	Phe	Ala	Pro	Ser	Ser	Asp	Ser	Ile	Leu	His	Ile	Arg	Glu	

ES 2 395 842 T3

290	295	300	
ccc aag gat gtc agt tca gtc gct gag aga atg aaa ttt ctc gct tca			960
Pro Lys Asp Val Ser Ser Val Ala Glu Arg Met Lys Phe Leu Ala Ser			
305	310	315	320
aat cca gaa gca tat aac caa tca ctg agg tgg aag ttt gag ggc cct			1008
Asn Pro Glu Ala Tyr Asn Gln Ser Leu Arg Trp Lys Phe Glu Gly Pro			
	325	330	335
tct aac tcc ttc aaa gcc ctg gtg gac atg gca gca gtt cac tcc tcc			1056
Ser Asn Ser Phe Lys Ala Leu Val Asp Met Ala Ala Val His Ser Ser			
	340	345	350
tgc cgc cta tgc att cac att gcc acc aag atc aga gag aag gaa gag			1104
Cys Arg Leu Cys Ile His Ile Ala Thr Lys Ile Arg Glu Lys Glu Glu			
	355	360	365
aga aac ccg aat ttc aag act cgc cct tgc aag tgc acc cgc aat ggg			1152
Arg Asn Pro Asn Phe Lys Thr Arg Pro Cys Lys Cys Thr Arg Asn Gly			
	370	375	380
tct acc tta tat cac tta tac gcc cgc gaa aga ggc acc ttt gac ttc			1200
Ser Thr Leu Tyr His Leu Tyr Ala Arg Glu Arg Gly Thr Phe Asp Phe			
	385	390	400
tta tca atc ttc atg aga tcg gat aat cta tca ctg aaa gcg ctg ggg			1248
Leu Ser Ile Phe Met Arg Ser Asp Asn Leu Ser Leu Lys Ala Leu Gly			
	405	410	415
tca aca gtt ctt gag aaa ttc agt tct ttg aag cac gtg ccg att tgg			1296
Ser Thr Val Leu Glu Lys Phe Ser Ser Leu Lys His Val Pro Ile Trp			
	420	425	430
aag aag gag agg cca gag agt ctg aaa gga ggg agc aag ctg gat ctt			1344
Lys Lys Glu Arg Pro Glu Ser Leu Lys Gly Gly Ser Lys Leu Asp Leu			
	435	440	445
tac aga atc tat cca gtg ggc att act cag aga gaa gct ctc ttc tct			1392
Tyr Arg Ile Tyr Pro Val Gly Ile Thr Gln Arg Glu Ala Leu Phe Ser			
	450	455	460
ttc cag ttc aac act gac aaa gaa ctt caa atc tac ctt gaa tcc cat			1440
Phe Gln Phe Asn Thr Asp Lys Glu Leu Gln Ile Tyr Leu Glu Ser His			
	465	470	475
cca tgt gcg aag ttt gaa gtc atc ttt att tga			1473
Pro Cys Ala Lys Phe Glu Val Ile Phe Ile *			
	485	490	

<210> 3

<211> 490

5 <212> PRT <

213> Lemna minor

<400> 3

ES 2 395 842 T3

Met	Ala	Thr	Ser	Ala	Ala	Gly	Ala	Leu	Asn	Ala	Gly	Gly	Arg	Val	Gly
1				5					10					15	
Gly	Arg	Arg	Ser	Trp	Val	Arg	Leu	Leu	Pro	Phe	Phe	Val	Leu	Met	Leu
			20					25					30		
Val	Val	Gly	Glu	Ile	Trp	Phe	Leu	Gly	Arg	Leu	Asp	Val	Val	Lys	Asn
		35					40					45			
Ala	Ala	Met	Val	Gln	Asn	Trp	Thr	Ser	Ser	His	Leu	Phe	Phe	Leu	Pro
	50					55					60				

ES 2 395 842 T3

Val	Ser	Ser	Tyr	Thr	Trp	Ser	Glu	Thr	Val	Lys	Glu	Glu	Glu	Asp	Cys	65	70	75	80
Lys	Asp	Trp	Leu	Glu	Arg	Val	Asp	Ala	Val	Asp	Tyr	Lys	Arg	Asp	Phe	85	90	95	
Arg	Val	Glu	Pro	Val	Leu	Val	Asn	Asp	Ala	Glu	Gln	Asp	Trp	Ser	Ser	100	105	110	
Cys	Ser	Val	Gly	Cys	Lys	Phe	Gly	Ser	Phe	Pro	Gly	Arg	Thr	Pro	Asp	115	120	125	
Ala	Thr	Phe	Gly	Phe	Ser	Gln	Asn	Pro	Ser	Thr	Val	Ser	Val	His	Arg	130	135	140	
Ser	Met	Glu	Ser	Ser	His	Tyr	Tyr	Leu	Glu	Asn	Asn	Leu	Asp	Asn	Ala	145	150	155	160
Arg	Arg	Lys	Gly	Tyr	Gln	Ile	Val	Met	Thr	Thr	Ser	Leu	Leu	Ser	Asp	165	170	175	
Val	Pro	Val	Gly	Tyr	Phe	Ser	Trp	Ala	Glu	Tyr	Asp	Ile	Met	Ala	Pro	180	185	190	
Leu	Gln	Pro	Lys	Thr	Ala	Gly	Ala	Leu	Ala	Ala	Ala	Phe	Ile	Ser	Asn	195	200	205	
Cys	Gly	Ala	Arg	Asn	Phe	Arg	Leu	Gln	Ala	Leu	Asp	Met	Leu	Glu	Lys	210	215	220	
Ser	Asn	Ile	Lys	Ile	Asp	Ser	Tyr	Gly	Ala	Cys	His	Arg	Asn	Gln	Asp	225	230	235	240
Gly	Lys	Val	Asp	Lys	Val	Gln	Thr	Leu	Lys	Arg	Tyr	Lys	Phe	Ser	Leu	245	250	255	
Ala	Phe	Glu	Asn	Ser	Asn	Glu	Asp	Asp	Tyr	Val	Thr	Glu	Lys	Phe	Phe	260	265	270	
Gln	Ser	Leu	Val	Ala	Gly	Ala	Ile	Pro	Val	Val	Val	Gly	Ala	Pro	Asn	275	280	285	
Ile	Gln	Asn	Phe	Ala	Pro	Ser	Ser	Asp	Ser	Ile	Leu	His	Ile	Arg	Glu	290	295	300	
Pro	Lys	Asp	Val	Ser	Ser	Val	Ala	Glu	Arg	Met	Lys	Phe	Leu	Ala	Ser	305	310	315	320
Asn	Pro	Glu	Ala	Tyr	Asn	Gln	Ser	Leu	Arg	Trp	Lys	Phe	Glu	Gly	Pro	325	330	335	
Ser	Asn	Ser	Phe	Lys	Ala	Leu	Val	Asp	Met	Ala	Ala	Val	His	Ser	Ser	340	345	350	
Cys	Arg	Leu	Cys	Ile	His	Ile	Ala	Thr	Lys	Ile	Arg	Glu	Lys	Glu	Glu	355	360	365	
Arg	Asn	Pro	Asn	Phe	Lys	Thr	Arg	Pro	Cys	Lys	Cys	Thr	Arg	Asn	Gly	370	375	380	
Ser	Thr	Leu	Tyr	His	Leu	Tyr	Ala	Arg	Glu	Arg	Gly	Thr	Phe	Asp	Phe	385	390	395	400
Leu	Ser	Ile	Phe	Met	Arg	Ser	Asp	Asn	Leu	Ser	Leu	Lys	Ala	Leu	Gly	405	410	415	
Ser	Thr	Val	Leu	Glu	Lys	Phe	Ser	Ser	Leu	Lys	His	Val	Pro	Ile	Trp	420	425	430	
Lys	Lys	Glu	Arg	Pro	Glu	Ser	Leu	Lys	Gly	Gly	Ser	Lys	Leu	Asp	Leu	435	440	445	
Tyr	Arg	Ile	Tyr	Pro	Val	Gly	Ile	Thr	Gln	Arg	Glu	Ala	Leu	Phe	Ser	450	455	460	
Phe	Gln	Phe	Asn	Thr	Asp	Lys	Glu	Leu	Gln	Ile	Tyr	Leu	Glu	Ser	His	465	470	475	480
Pro	Cys	Ala	Lys	Phe	Glu	Val	Ile	Phe	Ile							485	490		

<210> 4

<211> 1860

5 <212> ADN

<213> Lemna minor

<220>

ES 2 395 842 T3

<221> CDS

<222> (63)...(1592)

<400> 4

```

ggcttccaac cggaggatct cgagctgaag aatcttcatg actgaagaat tcatgtgatc 60
cc atg gct ttg gtg aac tcg cga ggg agc agg gtc aga cgc atc gcg 107
Met Ala Leu Val Asn Ser Arg Gly Ser Arg Val Arg Arg Ile Ala
1 5 10 15

aag ccc acc ttc gtt ttc ctc ttg atc aac gta gtc tgt ctc ctg tac 155
Lys Pro Thr Phe Val Phe Leu Leu Ile Asn Val Val Cys Leu Leu Tyr
20 25 30

ttt ttc cgt cag aac cct aat ccc att ccc gac gct tgt ctt cac ggg 203
Phe Phe Arg Gln Asn Pro Asn Pro Ile Pro Asp Ala Cys Leu His Gly
35 40 45

gaa tgc gac aaa ccc ccg att tta gtg act ccc cgg cga tgg aac ttg 251
Glu Cys Asp Lys Pro Pro Ile Leu Val Thr Pro Arg Arg Trp Asn Leu
50 55 60

aag cca tgg ccg att ctt cct tcc ttt ctg cca tgg gtg ccg agc tcc 299
Lys Pro Trp Pro Ile Leu Pro Ser Phe Leu Pro Trp Val Pro Ser Ser
65 70 75

cac cct gcc cag ggc tcc tgc gaa gcc tac ttc ggc aac agc ttc aac 347
His Pro Ala Gln Gly Ser Cys Glu Ala Tyr Phe Gly Asn Ser Phe Asn
80 85 90 95

cgc cgg acg gag atg ctg aag aag gta gag gga aga gga tgg ttc cag 395
Arg Arg Thr Glu Met Leu Lys Lys Val Glu Gly Arg Gly Trp Phe Gln
100 105 110

tgc ctg tac agc gat act ctt cga agt tct gtt tgc cag gga ggg aat 443
Cys Leu Tyr Ser Asp Thr Leu Arg Ser Ser Val Cys Gln Gly Gly Asn
115 120 125

ttg cgg atg gac ccg gaa agg att agg atg tcg aaa ggg ggg gaa gat 491
Leu Arg Met Asp Pro Glu Arg Ile Arg Met Ser Lys Gly Gly Glu Asp
130 135 140

cta gag gag gtg atg aag aga gag gag gaa gaa gaa ttg ccc aaa ttc 539
Leu Glu Glu Val Met Lys Arg Glu Glu Glu Glu Glu Leu Pro Lys Phe
145 150 155

gag gag ggg tcg ttc cag att gaa tct ggt tat gga agc gga ggg gaa 587
Glu Glu Gly Ser Phe Gln Ile Glu Ser Gly Tyr Gly Ser Gly Gly Glu
160 165 170 175

gtt gga gag aga att gcg act gac gag gtc ctc gat aat gtt gtg ccg 635
Val Gly Glu Arg Ile Ala Thr Asp Glu Val Leu Asp Asn Val Val Pro
180 185 190

aaa ggc gct gtt cat gta cat acc atg cgc aat ctc atc agt tcg att 683
Lys Gly Ala Val His Val His Thr Met Arg Asn Leu Ile Ser Ser Ile
195 200 205

cag att gtt ggt ccc ggg cat ctt caa tgc tct cag tgg atc gac gaa 731
Gln Ile Val Gly Pro Gly His Leu Gln Cys Ser Gln Trp Ile Asp Glu
210 215 220

ccg gtt ctt ctt gtc aca cgc ttc gaa tac gcc aat ctc ttt cac acc 779
Pro Val Leu Leu Val Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr
225 230 235

gtc acc gac tgg tac agc gcc tac gca agc tcg agg att gcc aac ttg 827
Val Thr Asp Trp Tyr Ser Ala Tyr Ala Ser Ser Arg Ile Ala Asn Leu
240 245 250 255

```


ES 2 395 842 T3

cct tct cgc cct cac tta att ttc gtc gat ggc cat tgc agg gcg gaa	875
Pro Ser Arg Pro His Leu Ile Phe Val Asp Gly His Cys Arg Ala Glu	
260 265 270	
cag tta gag gac atg tgg aga gcc ctg ttc tcg acc gtc cga tac tcc	923
Gln Leu Glu Asp Met Trp Arg Ala Leu Phe Ser Thr Val Arg Tyr Ser	
275 280 285	
aag aac ttc tcc cag cca atc tgc ttc cgc cac gtc gtc ctc tca cct	971
Lys Asn Phe Ser Gln Pro Ile Cys Phe Arg His Val Val Leu Ser Pro	
290 295 300	
ctg ggc tat gag acg gct ctc ttc aaa ggc cta tca gag agc ttc agc	1019
Leu Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Ser Phe Ser	
305 310 315	
tgt gag gga gct cgg gcc aat cgg ctc aaa gtc aac ccc gat gac cag	1067
Cys Glu Gly Ala Pro Ala Asn Arg Leu Lys Val Asn Pro Asp Asp Gln	
320 325 330 335	
aag act gca aga ctg gct gaa ttc gga gag atg atc aga gcc gcc ttt	1115
Lys Thr Ala Arg Leu Ala Glu Phe Gly Glu Met Ile Arg Ala Ala Phe	
340 345 350	
gac ttt cct gtc gtt gac cgg tcc att gac cgg ttg acc aaa tcc atc	1163
Asp Phe Pro Val Val Asp Pro Ser Ile Asp Pro Leu Thr Lys Ser Ile	
355 360 365	
ctc ttc gtg cgg cgg gaa gat tac gtg gcg cac cca cgc cac agt ggg	1211
Leu Phe Val Arg Arg Glu Asp Tyr Val Ala His Pro Arg His Ser Gly	
370 375 380	
aga gtg gag tcg cgg ctg acc aac gag caa gag gtg ttt gac ttt ctg	1259
Arg Val Glu Ser Arg Leu Thr Asn Glu Gln Glu Val Phe Asp Phe Leu	
385 390 395	
cac aat tgg gca agt cat cac aga ggc agg tgc aac atc agt atg gtc	1307
His Asn Trp Ala Ser His His Arg Gly Arg Cys Asn Ile Ser Met Val	
400 405 410 415	
aac ggg ctt ttc gcg cac atg gga atg aag gaa cag ttg aag gcg att	1355
Asn Gly Leu Phe Ala His Met Gly Met Lys Glu Gln Leu Lys Ala Ile	
420 425 430	
atg gaa gct tcg gtg gtg gtg ggg gcc cac ggg gct ggt ttg acc cat	1403
Met Glu Ala Ser Val Val Val Gly Ala His Gly Ala Gly Leu Thr His	
435 440 445	
ctg gtg gca gca agg tca acg aca gtt gtt ctt gag att ctg agt agt	1451
Leu Val Ala Ala Arg Ser Thr Thr Val Val Leu Glu Ile Leu Ser Ser	
450 455 460	
caa tac cgt aga ccg cac ttt caa ctg att tct cgg tgg aaa ggg ttg	1499
Gln Tyr Arg Arg Pro His Phe Gln Leu Ile Ser Arg Trp Lys Gly Leu	
465 470 475	
gac tac cat gca att aat ctt gcc ggg tca ttt gct gac cct cgg gag	1547
Asp Tyr His Ala Ile Asn Leu Ala Gly Ser Phe Ala Asp Pro Arg Glu	
480 485 490 495	
gtg gtc gag aaa ttg act ggc ata gtt gac agg ctt gga tgt tga	1592
Val Val Glu Lys Leu Thr Gly Ile Val Asp Arg Leu Gly Cys *	
500 505	
agagaagtga aagtcaacat ttggaatttt aactttaagg ggtggttaac aattgagcgg	1652
cattgtcaac gggtttggat gctgggaaaa gtgaaaatca acacttggag ttctgacatt	1712
gaaggcaaga cgtggaattt tgatggtgtt gaggatattt ggatgtggag ttctgatgaa	1772
ttaaagcagg ggttgatcat ttgccagtgg aattatgttg gtgtaagaga gaagggggag	1832
aataaacagt gttagagagc tatgctgg	1860

ES 2 395 842 T3

<210> 5
<211> 1530
<212> ADN
<213> Lemna minor

5

<220>
<221> CDS
<222> (1)...(1530)
<223> XylT isoforma #1; Codifica beta-1, 2-xilosiltransferasa

10

<400> 5

atg gct ttg gtg aac tcg cga ggg agc agg gtc aga cgc atc gcg aag	48
Met Ala Leu Val Asn Ser Arg Gly Ser Arg Val Arg Arg Ile Ala Lys	
1 5 10 15	
ccc acc ttc gtt ttc ctc ttg atc aac gta gtc tgt ctc ctg tac ttt	96
Pro Thr Phe Val Phe Leu Leu Ile Asn Val Val Cys Leu Leu Tyr Phe	
20 25 30	
ttc cgt cag aac cct aat ccc att ccc gac gct tgt ctt cac ggg gaa	144
Phe Arg Gln Asn Pro Asn Pro Ile Pro Asp Ala Cys Leu His Gly Glu	
35 40 45	
tgc gac aaa ccc ccg att tta gtg act ccc cgg cga tgg aac ttg aag	192
Cys Asp Lys Pro Pro Ile Leu Val Thr Pro Arg Arg Trp Asn Leu Lys	
50 55 60	
cca tgg ccg att ctt cct tcc ttt ctg cca tgg gtg ccg agc tcc cac	240
Pro Trp Pro Ile Leu Pro Ser Phe Leu Pro Trp Val Pro Ser Ser His	
65 70 75 80	
cct gcc cag ggc tcc tgc gaa gcc tac ttc ggc aac agc ttc aac cgc	288
Pro Ala Gln Gly Ser Cys Glu Ala Tyr Phe Gly Asn Ser Phe Asn Arg	
85 90 95	
cgg acg gag atg ctg aag aag gta gag gga aga gga tgg ttc cag tgc	336
Arg Thr Glu Met Leu Lys Lys Val Glu Gly Arg Gly Trp Phe Gln Cys	
100 105 110	
ctg tac agc gat act ctt cga agt tct gtt tgc cag gga ggg aat ttg	384
Leu Tyr Ser Asp Thr Leu Arg Ser Ser Val Cys Gln Gly Gly Asn Leu	
115 120 125	
cgg atg gac ccg gaa agg att agg atg tcg aaa ggg ggg gaa gat cta	432
Arg Met Asp Pro Glu Arg Ile Arg Met Ser Lys Gly Gly Glu Asp Leu	
130 135 140	
gag gag gtg atg aag aga gag gag gaa gaa gaa ttg ccc aaa ttc gag	480
Glu Glu Val Met Lys Arg Glu Glu Glu Glu Glu Leu Pro Lys Phe Glu	
145 150 155 160	
gag ggg tcg ttc cag att gaa tct ggt tat gga agc gga ggg gaa gtt	528
Glu Gly Ser Phe Gln Ile Glu Ser Gly Tyr Gly Ser Gly Gly Glu Val	
165 170 175	
gga gag aga att gcg act gac gag gtc ctc gat aat gtt gtg ccg aaa	576
Gly Glu Arg Ile Ala Thr Asp Glu Val Leu Asp Asn Val Val Pro Lys	
180 185 190	

ES 2 395 842 T3

ggc gct gtt cat gta cat acc atg cgc aat ctc atc agt tgc att cag Gly Ala Val His Val His Thr Met Arg Asn Leu Ile Ser Ser Ile Gln 195 200 205	624
att gtt ggt ccc ggg cat ctt caa tgc tct cag tgg atc gac gaa ccg Ile Val Gly Pro Gly His Leu Gln Cys Ser Gln Trp Ile Asp Glu Pro 210 215 220	672
gtt ctt ctt gtc aca cgc ttc gaa tac gcc aat ctc ttt cac acc gtc Val Leu Leu Val Thr Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val 225 230 235 240	720
acc gac tgg tac agc gcc tac gca agc tgc agg att gcc aac ttg cct Thr Asp Trp Tyr Ser Ala Tyr Ala Ser Ser Arg Ile Ala Asn Leu Pro 245 250 255	768
tct cgc cct cac tta att ttc gtc gat gcc cat tgc agg gcg gaa cag Ser Arg Pro His Leu Ile Phe Val Asp Gly His Cys Arg Ala Glu Gln 260 265 270	816
tta gag gac atg tgg aga gcc ctg ttc tgc acc gtc cga tac tcc aag Leu Glu Asp Met Trp Arg Ala Leu Phe Ser Thr Val Arg Tyr Ser Lys 275 280 285	864
aac ttc tcc cag cca atc tgc ttc cgc cac gtc gtc ctc tca cct ctg Asn Phe Ser Gln Pro Ile Cys Phe Arg His Val Val Leu Ser Pro Leu 290 295 300	912
ggc tat gag acg gct ctc ttc aaa gcc cta tca gag agc ttc agc tgt Gly Tyr Glu Thr Ala Leu Phe Lys Gly Leu Ser Glu Ser Phe Ser Cys 305 310 315 320	960
gag gga gct ccg gcc aat cgg ctc aaa gtc aac ccc gat gac cag aag Glu Gly Ala Pro Ala Asn Arg Leu Lys Val Asn Pro Asp Asp Gln Lys 325 330 335	1008
act gca aga ctg gct gaa ttc gga gag atg atc aga gcc gcc ttt gac Thr Ala Arg Leu Ala Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Asp 340 345 350	1056
ttt cct gtc gtt gac ccg tcc att gac ccg ttg acc aaa tcc atc ctc Phe Pro Val Val Asp Pro Ser Ile Asp Pro Leu Thr Lys Ser Ile Leu 355 360 365	1104
ttc gtg cgg cgg gaa gat tac gtg gcg cac cca cgc cac agt ggg aga Phe Val Arg Arg Glu Asp Tyr Val Ala His Pro Arg His Ser Gly Arg 370 375 380	1152
gtg gag tgc cgg ctg acc aac gag caa gag gtg ttt gac ttt ctg cac Val Glu Ser Arg Leu Thr Asn Glu Gln Glu Val Phe Asp Phe Leu His 385 390 395 400	1200
aat tgg gca agt cat cac aga gcc agg tgc aac atc agt atg gtc aac Asn Trp Ala Ser His His Arg Gly Arg Cys Asn Ile Ser Met Val Asn 405 410 415	1248
ggg ctt ttc gcg cac atg gga atg aag gaa cag ttg aag gcg att atg Gly Leu Phe Ala His Met Gly Met Lys Glu Gln Leu Lys Ala Ile Met 420 425 430	1296
gaa gct tgc gtg gtg gtg ggg gcc cac ggg gct ggt ttg acc cat ctg Glu Ala Ser Val Val Val Gly Ala His Gly Ala Gly Leu Thr His Leu 435 440 445	1344

ES 2 395 842 T3

gtg gca gca agg tca acg aca gtt gtt ctt gag att ctg agt agt caa	1392
Val Ala Ala Arg Ser Thr Thr Val Val Leu Glu Ile Leu Ser Ser Gln	
450 455 460	
tac cgt aga ccg cac ttt caa ctg att tct cgg tgg aaa ggg ttg gac	1440
Tyr Arg Arg Pro His Phe Gln Leu Ile Ser Arg Trp Lys Gly Leu Asp	
465 470 475 480	
tac cat gca att aat ctt gcc ggg tca ttt gct gac cct cgg gag gtg	1488
Tyr His Ala Ile Asn Leu Ala Gly Ser Phe Ala Asp Pro Arg Glu Val	
485 490 495	
gtc gag aaa ttg act ggc ata gtt gac agg ctt gga tgt tga	1530
Val Glu Lys Leu Thr Gly Ile Val Asp Arg Leu Gly Cys *	
500 505	

<210> 6

<211> 509

5 <212> PRT

<213> Lemna minor

<400> 6

Met	Ala	Leu	Val	Asn	Ser	Arg	Gly	Ser	Arg	Val	Arg	Arg	Ile	Ala	Lys
1				5					10					15	
Pro	Thr	Phe	Val	Phe	Leu	Leu	Ile	Asn	Val	Val	Cys	Leu	Leu	Tyr	Phe
			20					25					30		
Phe	Arg	Gln	Asn	Pro	Asn	Pro	Ile	Pro	Asp	Ala	Cys	Leu	His	Gly	Glu
			35				40					45			
Cys	Asp	Lys	Pro	Pro	Ile	Leu	Val	Thr	Pro	Arg	Arg	Trp	Asn	Leu	Lys
	50					55					60				
Pro	Trp	Pro	Ile	Leu	Pro	Ser	Phe	Leu	Pro	Trp	Val	Pro	Ser	Ser	His
65					70					75					80
Pro	Ala	Gln	Gly	Ser	Cys	Glu	Ala	Tyr	Phe	Gly	Asn	Ser	Phe	Asn	Arg
				85					90					95	
Arg	Thr	Glu	Met	Leu	Lys	Lys	Val	Glu	Gly	Arg	Gly	Trp	Phe	Gln	Cys
			100					105					110		
Leu	Tyr	Ser	Asp	Thr	Leu	Arg	Ser	Ser	Val	Cys	Gln	Gly	Gly	Asn	Leu
			115				120					125			
Arg	Met	Asp	Pro	Glu	Arg	Ile	Arg	Met	Ser	Lys	Gly	Gly	Glu	Asp	Leu
	130					135					140				
Glu	Glu	Val	Met	Lys	Arg	Glu	Glu	Glu	Glu	Glu	Leu	Pro	Lys	Phe	Glu
145					150					155					160
Glu	Gly	Ser	Phe	Gln	Ile	Glu	Ser	Gly	Tyr	Gly	Ser	Gly	Gly	Glu	Val
				165					170					175	
Gly	Glu	Arg	Ile	Ala	Thr	Asp	Glu	Val	Leu	Asp	Asn	Val	Val	Pro	Lys
			180					185					190		
Gly	Ala	Val	His	Val	His	Thr	Met	Arg	Asn	Leu	Ile	Ser	Ser	Ile	Gln
		195				200						205			
Ile	Val	Gly	Pro	Gly	His	Leu	Gln	Cys	Ser	Gln	Trp	Ile	Asp	Glu	Pro
	210					215					220				
Val	Leu	Leu	Val	Thr	Arg	Phe	Glu	Tyr	Ala	Asn	Leu	Phe	His	Thr	Val
225					230					235					240
Thr	Asp	Trp	Tyr	Ser	Ala	Tyr	Ala	Ser	Ser	Arg	Ile	Ala	Asn	Leu	Pro
				245					250					255	
Ser	Arg	Pro	His	Leu	Ile	Phe	Val	Asp	Gly	His	Cys	Arg	Ala	Glu	Gln
			260					265					270		
Leu	Glu	Asp	Met	Trp	Arg	Ala	Leu	Phe	Ser	Thr	Val	Arg	Tyr	Ser	Lys
		275					280					285			
Asn	Phe	Ser	Gln	Pro	Ile	Cys	Phe	Arg	His	Val	Val	Leu	Ser	Pro	Leu
	290					295					300				
Gly	Tyr	Glu	Thr	Ala	Leu	Phe	Lys	Gly	Leu	Ser	Glu	Ser	Phe	Ser	Cys
305					310					315					320
Glu	Gly	Ala	Pro	Ala	Asn	Arg	Leu	Lys	Val	Asn	Pro	Asp	Asp	Gln	Lys

ES 2 395 842 T3

Thr	Ala	Arg	Leu	Ala	Glu	Phe	Gly	Glu	Met	Ile	Arg	Ala	Ala	Phe	Asp
			340					345					350		
Phe	Pro	Val	Val	Asp	Pro	Ser	Ile	Asp	Pro	Leu	Thr	Lys	Ser	Ile	Leu
		355					360					365			
Phe	Val	Arg	Arg	Glu	Asp	Tyr	Val	Ala	His	Pro	Arg	His	Ser	Gly	Arg
		370				375					380				
Val	Glu	Ser	Arg	Leu	Thr	Asn	Glu	Gln	Glu	Val	Phe	Asp	Phe	Leu	His
385					390					395					400
Asn	Trp	Ala	Ser	His	His	Arg	Gly	Arg	Cys	Asn	Ile	Ser	Met	Val	Asn
				405					410					415	
Gly	Leu	Phe	Ala	His	Met	Gly	Met	Lys	Glu	Gln	Leu	Lys	Ala	Ile	Met
			420					425					430		
Glu	Ala	Ser	Val	Val	Val	Gly	Ala	His	Gly	Ala	Gly	Leu	Thr	His	Leu
		435					440					445			
Val	Ala	Ala	Arg	Ser	Thr	Thr	Val	Val	Leu	Glu	Ile	Leu	Ser	Ser	Gln
		450				455					460				
Tyr	Arg	Arg	Pro	His	Phe	Gln	Leu	Ile	Ser	Arg	Trp	Lys	Gly	Leu	Asp
465					470					475					480
Tyr	His	Ala	Ile	Asn	Leu	Ala	Gly	Ser	Phe	Ala	Asp	Pro	Arg	Glu	Val
				485					490					495	
Val	Glu	Lys	Leu	Thr	Gly	Ile	Val	Asp	Arg	Leu	Gly	Cys			
			500					505							

<210> 7

<211> 2160

5 <212> ADN

<213> Lemna minor

<400> 7

ES 2 395 842 T3

```

agtcgagtga tatgaaatct tgggtgaagaa ggatcggaga acggaccggg tgaggcaagg 60
ataattctgc tgttaaattc gagagcaaga cacctgcaat tcaagaatcg agtggcaatt 120
aatatagcag gatgatctgg aaggtagatc ctgcccacg aatgatccaa acatcaacac 180
taggatcata ccgttaacaa taatgaatga aaaagtagaa gatgacgaag ttgaagtgat 240
gaccaaaaaa tttgaaaatt ccaaccgtat ggccggaaac agtgtgaaga aaatcgaat 300
caaatactct aatggatcgg attgttattc tggaggcaaa tctgaaactt cgaggatagg 360
atttaatcca cgcaagtaat aatttgaaac tcagaaggag aaaaaaaac taaaatagag 420
aagaagagat ctcaagaag ccgtgagcac gagacgaacg agaagaggta aagcaccagt 480
cagaggaaaa caccaaaatt agagaaatag cacgaacatt aaagcacaga tccgcgccgc 540
aaacccgaaa gacgaaaaat agagccaaac gaaaccctaa taatcgatct gcacaaaaaa 600
aaaaaaaaaa aactttgaga agagccgcga aattacccta gaatcctcag aactggccgg 660
acgagagaag cgctcgatcg aaacccaaca taaaaccctt tccaacggca aattactccg 720
caaaacccga aaaataaaca aaatcaacga tcacgagaag gtgcaagggc aaaaagggc 780
agtgcgatcg agagtctacc tgaatcgctg gcgcaaaagg cgagcccacc gacgaacgct 840
ccctctagaa cctggagatg cggcgagaga gaaggaaaga tcttcgggtg gtgatgctcg 900
ctatttatcg caagagagtt agagagatct tcttcggcgg cggattttctg gcctctagcg 960
tttaacctca ccgccagtg ctacatcct tcttctcata tttgaatatt taattaacaa 1020
atgaatcagt catttttctt taatttttaa ttcccgaga gggcaatgtt ggtatcaaaa 1080
attatttagg aaaaattaat tacacgaata atcggatttt tccctttttt taattaattt 1140
ctaattttgg aaaaggaaag aaaaatttta ggggtatgga gggcaagaat gaaatattac 1200
aaaattaggg gtttttgcgt aattttattat atttaataaa gaaagtcgaa tattcccatc 1260
cgattggtag ttgaaagggg ccgaaaggcc tcgggggttc tagagatttc tacattatc 1320
tcgtttttgt cgccaagaag gtgggcaatt atgtttcatg ccttaacttc ttctttttgt 1380
gggaatactc ttattcttag tacaaaagaa aagagtatat gcataataaa gatgaaaaat 1440
gggtttatcc gagatttcta cgtcatgtgt gactcgctta ggaaatatcg ccgaaacctt 1500
acaaaggcgg tacgctcctc tccccgcacc tataaataga gacctttgcc tcgtctttct 1560
caactcaagc atttctgtat gatccttctc tttccgcgga agctctcgcg ccagttgatc 1620
gcaaggtagt cgtctttcct ccttgtgatt cgatctttct gttggctaga tctggcttat 1680
tgatctgctc tattgatctg gtctatttat cgctgcacg ggatctattg atccgtatgt 1740
tgatttggga tccgtaggtt ggtttggatc ggagactgcg atttgattct tgtgatttcg 1800
cttggatttc ggaaatcggg gtggttgaaag tcgtgcgac ttttagatct gctccttttt 1860
ttatttgcta ttttatattt acgttggtta tgatcgcgga ttattttgat tcgtttattc 1920
gagatccatg ccgtttaact cgttctttgt gctccgatct ttgcgatacg tcggtcgttc 1980
tagatccgtt cactaggtta gttttaagtt ctttgagctt gatttatatg gatttgctgt 2040

```

```

tttccaggaa aaatttatgc gcgattctta cgcccgtttc cccattttac tttaggtcgt 2100
gaattctttt gatctgagaa tgatgaatct gacatgtacc ttccggtttg taatttgacg 2160

```

<210> 8

<211> 2021

5 <212> ADN

<213> Spirodela polyrrhiza

<400> 8

```

caaataaaga gatggacaga taatgagatg aattagaaaa aaaaaattcg tgttgtaaga 60
tagaataactt gctatctact gatgaatgca gttcagtttt cctcacgatac ttaaagatcg 120
cgactatcc tcagcttcac tctggaaatt ttgattctct tcttctgctc agcagcctcg 180
actctgtcta gggtttcgta caatcggacg ccattctaca tgaatcgagc acagggaatg 240
aagacaatta ggagatcctc gatgtcctcc gacttacttg catgacttga cggggaagat 300
ctcgagcagg gaagcgacgc ctctccggag gactcgctc gccgagagga cctcctccgc 360
gacacggacc atggcctcca cggggtagaa gctggccctg ttctttattc tcttgaggat 420
catcggccga agcctccgca aatccatccc cgaggagtag aatctcgctc gcaggaagca 480
tctgtcgaga tcctcgccga ggccggcgag atacctcgcc ggccgccgca tggcgccggg 540
gacggagcac caccacggag aagaagaacc ctaaccaag gcattaacga agttgcgcag 600
attatacaaaa agcctcaca tatctttcat tttctatttc actgatacat tttcattatt 660
gtatatgagt gtttatttaa attattccgt attagaaaaa cactccaga acccgacaaa 720
atagggtgac gtcacatgag tgatcatgac gcccaacagc cgcagattta aaatcgggtg 780
atgagtgcgg ccacgcccag aaagcgatgg gccttcgtcg atgccgtgag aatccatctg 840
acataaagta aacggcgccg tcagtattga cggcgtaga cactggaaa gaagctattg 900
gttcacgcat cgggtggtcc gctagcctcc gtcgaccgct agtactataa atacggtccc 960
gaggcctcct caccactcgc acatatcctc tttgttttcc tctccgtgaa agaagcgagg 1020
aagcgcgctc tctctcccaa ggtaaggagc agatctcttt gatcggtttt gttcttcttt 1080
tggtttgttt ttttttctg cggatcttcg gttgcatcat gccttggtcg tttttattag 1140
tttaggatat cctcgtttg atctgagccg atcatatatg ttaaagggtg tgttcgatct 1200
ctttgttcat ttccgcatga aaaggatgta tccttttgat gtgaggcgat cttctatggg 1260
taagactttg ttccgtctat tgatcatttc tgttcttcgt ttttgagttt tttctgcgg 1320
atategcatc atccctaggt ttttgctttg gttaggatgc atcctttgga tttgagccga 1380
tctcccttg ttaaggctgt gtctgttgca gaggagaaag totgtcgagg tccttatgca 1440
ggctttgtcc agatgcgcgt gctctctcat gctatgaatt tatgttttga gaactcctcc 1500
cgggttttct agatccggat ttgaagtatt cattgcggtt ccccttcggt tttatgtatt 1560
tctcgagttg atttggtcca tgatcgtgtt ctgtccagat ctctcttgat atggatgaga 1620
tattcgttac ctctttcaaa catcgggtga tgttcttttt agtcttggtc cacctttatc 1680
tagaaattaa ttttcggtt gaaacccctg cttgttaagg tgatgtattc cttctttata 1740
gatttcgggt tggtatttct taacggtgat ctgtccgata catgtgttgc acctctgtt 1800
ttctgtgtaa tcctctgtga attataatta tgttttgaaa acgtacttaa gtaaggggca 1860
tgttcccggt ttaaaacttt tgttctatca atttgtggtt aatagatcct gatttggtgg 1920
cgccttatcc tgtctttaat cgtggatttt atttatcttg agcgcgtcct tttcttttaa 1980
aatcatgtgt ttaacctttc agtcgtcata tgttccatca g 2021

```

<210> 9

<211> 2068

5 <212> ADN

<213> Lemna aequinoctialis

<400> 9

```

agtgtacca ttttttaaac cctacattta tcattcttta ttcattattg ccataagtta 60
atgaatattg aaattcaaat acgcgcaaga tgtcaatata gatcgaatat gaataccaga 120
tataaaatca aaaatcaaat atcaaaatta taaagatata aaatattgaa tccaaaagca 180
ataaagaata tcactattaa tatcaaaata tcgatttgaa gttcaaaaat tgggtccatt 240
aggagccaag accgatcatg atccgatact gatatacata tctgtagctc agtggttagg 300
cccctcaatt tgcttgccg aaggcagtg acaaaacctg gctctcgcaa gggcaaagaa 360
agagtctttc ccaaaaaaaa aaaaatcgaa cccatttgta gtatccaata ttgggattga 420
cataagatac caaaacataa agtactaacc acccaatott ataattaatc aagatttata 480
tcacatccaa tatcaagatc cgatatcaat acctagaccg gtaaacctta atttactott 540
ccccctcta aaaatttcca ataaatatct ccacatattt aactattaaa aaattgataa 600
gagataggcc ctagccctaa gtcctaacat ataaccactc tctatgaaaa gtcctattaa 660
atgacgtcat ttatttattt attgccggtt ggctgtcca cagccgcaat ttaatggatg 720
gctgacacgg caggaaccg acggcggtg ccgtgggaat aattctagag taaacctaac 780
ggcgccggtta actttgacgg tggcgaagac gcgtggggat aggtggttgg tccgcgtgac 840

```



```

ggcggcggtt cagcccgctc accttgagcc gagactataa atcgaggcga agggatgagc 900
tttgccattg cgttcttctt ctgttcatct ctgaaattcg ggcggaatcc ttcttcttct 960
caagggtatgg gctcgcgtct ttctgtttca atcgagtttt gatcttcggt ttggcggcga 1020
tcgggtgtttt ctttgtattg tgaataaatc cttgataaga aaacctagg ttttgtgacc 1080
tggtgacgga tgcgtgcgga tctgttattt gtcttttagg cgattttctc ttgtttgtaa 1140
tagtttatca taaccagatg aacatggatc aagtcgattt gacttatttt ttctgtgaaa 1200
ttaggccgaa atcctttttt ttggtttgag ccttgatatt tctatataat tcgatttgat 1260
tttttgtttt cttctgcgtc tgatgcttcc tcttgactcc tgattaaatt tttgctacgg 1320
aaacctaga tgcgcgagtc tgttgacaga ttctggcaaa tctgttttta tcataatcag 1380
atgaacgcaa attaatgcca ttggtttttt ctctgaaatt aggggggaaa ctccttatag 1440
tatgagcctc gataattcta taatagtcga tttgattttc tcttgccctc tgattcaatt 1500
tttgggtgcgg aaacctaga tattgtaatc tgtttacgga tgcctgcgga tctgattttt 1560
aatattgtga tctattgacg gatgctcgta gatctggttg ttttgatttc ttcatgcctt 1620
atacggcgat ttgattcggc gattaaaaat tttcaattct tttaaaaaaa atattaagat 1680
tttcaacggt tcaaatattt tcatagatcg gcacaaatac ttttcatcag attcctcctg 1740
atgtgatggg ttgtgtttta aatctgttga agatatcaga ttctattagg tcaccgatat 1800
aatcttctct gtttattctg cgatcggtgc ttacaaaccc tatttcctac ggtgattaat 1860
tatttttaat ctcttagcta gcgtaaatat atattttttt aatttgatct ttgcattagt 1920
ttctcctctt tctttgctat taattgtaac agatgctaca aaacatcaga ttttttttcc 1980
caattcggtt tcatcattat agaaaacttt tatctgatat ttttaatcgt cattaatata 2040
attttcaatt tattattttc ccttgacg          2068

```

<210> 10

<211> 1625

5 <212> ADN

<213> Lemna minor

<400> 10

```

agtcgagtgat tatgaaatct tgggtgaagaa ggatcggaga acggaccggg tgaggcaagg 60
ataattctgc tgttaaatcc gagagcaaga cacctgcaat tcaagaatcg agtggcaatt 120
aatatagcag gatgatctgg aaggtagatc ctgccatcg aatgatccaa acatcaacac 180
taggatcata ccgttaacaa taatgaatga aaaagtagaa gatgacgaag ttgaagtgat 240
gacaaaaaac tttgaaaatt ccaaccgtat ggccgggaatc agtgtgaaga aaatcgaaat 300
caaatactct aatggatcgg attgttatcc tggaggcaaa tctgaaactt cgaggatagg 360
atthaatcca cgcaagtaat aatttgaaac tcagaaggag aaaaaaaac taaaatagag 420
aagaagagat ctcaaagaag ccgtgagcac gagacgaacg agaagaggta aagcaccagt 480
cagaggaaaa caccaaaatt agagaaatag cacgaacatt aaagcacaga tccgcgccgc 540
aaaccgaaa gacgaaaaat agagccaaac gaaaccctaa taatcgatct gcacaaaaaa 600
aaaaaaaaaa aactttgaga agagccgcga aattacccta gaatcctcag aactggccgg 660
acgagagaag cgctcgatcg aaacccaaca taaacccctc tccaacggca aattactcgg 720
caaaacccga aaaataaaca aatcaacga tcacgagaag gtgcaagggc aaaaagaggc 780
agtgcgatcg agagtctacc tgaatcgtcg gcgcaaaagg cgagcccacc gacgaacgct 840
cctctagaa cctggagatg cggcgagaga gaaggaaaaga tcttcggtgg gtgatgctcg 900
ctatttatcg caagagagtt agagagatct tcttcggcgg cggatttctg gcatctagcg 960
tttaacctca ccgccagtg ctacatcct tcttctcata tttgaatatt taattaacaa 1020
atgaatcagt catttttctt taatttttaa ttcccgaga gggcaatgtt ggtatcaaaa 1080
attatttagg aaaaattaat tacacgaata atcgattttt tccctttttt taattaattt 1140
ctaattttgg aaaaggaaa gaaaatttta ggggtatgga gggcaagaat gaaatattac 1200
aaaattaggg gtttttgcgt aatttattat atttaataaa gaaagtcgaa tattcccatc 1260
cgatttgtag ttgaaagggg ccgaaaggcc tcgggggttc tagagatttc tacattattc 1320
tcgtttttgt cgccaaagag gtgggcaatt atgtttcatg ccttaacttc ttctttttgt 1380
gggaatactc ttattcttag tacaaaagaa aagagtatat gcataaataa gatgaaaaat 1440
gggtttattc gagatttcta cgtcatgtgt gactcgctta ggaaatatcg ccgaaacctc 1500
acaaaggcgg tacgctctc tcccccgacc tataaataga gacctttgcc tcgtctttct 1560
caactcaagc atttctgtat gatccttctc tttccgcgga agctctcgcg ccagttgatc 1620
gcaag          1625

```

10

<210> 11

<211> 1041
 <212> ADN
 <213> Spirodela polyrrhiza

5 <400> 11

caaataaaga gatggacaga taatgagatg aattagaaaa aaaaaattcg tgttgtaaga 60
 tagaataactt gctatctact gatgaatgca gttcagtttt cctcacgatac ttaaagatcg 120

cgactatccc tcagcttcac tctggaaatt ttgattctct tcttctgctc agcagcctcg 180
 actctgtcta ggggtttcgt caatcggacg ccattctaca tgaatcgagc acaggggaatg 240
 aagacaatta ggagatcctc gatgtcctcc gacttacttg catgacttga cggggaagat 300
 ctcgagcagg gaagcgacgc ctctccggag gactcgctc gccgagagga cctcctccgc 360
 gacacggacc atggcctcca cggggtagaa gctggcctg ttctttattc tcttgaggat 420
 catcggccga agcctccgca aatccatccc cgaggagtag aatctcgctc gcaggaagca 480
 tctgtcgaga tctcgcgca ggcggcggag atacctcgcc ggcgcgcgca tggcgccggg 540
 gacggagcac caccacggag aagaagaacc ctaaccacag gcattaacga agttgcgcag 600
 attatacaaa agcctcaca tatctttcat tttctatttc actgatacat tttcattatt 660
 gtatatgagt gtttatttaa attattccgt attagaaaag cacctccaga acccgacaaa 720
 atagggtgac gtcacatggt tgtcatgacc gcccaacagc cgcagattta aaatcggtg 780
 atgagtgcgg ccacgccacg aaagcagtg gcttctgtcg atgccgtgag aatccatctg 840
 acataaagta aacggcgccg tcagtattga cggcgtatga cacgtggaaa gaagctattg 900
 gttcacgcat cgggtggttc gctagcctcc gtcgaccgct agtactataa atacggtccc 960
 gaggcctcct caccactcgc acatatectc tttgttttcc tctcgtgaa agaagcgagg 1020
 aagcgcgtcg tctctcccaa g 1041

<210> 12
 <211> 964
 <212> ADN
 <213> Lemna aequinoctialis

<400> 12

agtgtaccaaa ttttttaaac cctacattta tcattcttta ttcattattg ccataagtta 60
 atgaatattg aaattcaaat acgcgcaaga tgtcaatata gatcgaatat gaataccaga 120
 tataaaatca aaaatcaaat atcaaatata taaagatata aaatattgaa tccaaaagca 180
 ataaagaata tcaactattaa tatcaaaata tcgatttgaa gttcaaaaat tgggtccatt 240
 aggagccaag accgatcatg atccgatact gatataata tctgtagctc agtggctagg 300
 cccctcaatt tgccctggccg aaggcagtg acaaaacctg gctctcgcaa gggcaaagaa 360
 agagtctttc ccaaaaaaaa aaaaatcgaa cccatttgta gtatccaata tttggattga 420
 cataagatac caaaacataa agtactaacc acccaatctt ataattaatc aagatttata 480
 tcacatccaa tatcaagatc cgatatcaat acctagaccg gtaaacccta atttactctt 540
 ccccccctcta aaaatttcca ataaatatct ccacatattt aactattaaa aaattgataa 600
 gagataggcc ctagccctaa gtcctaacat ataaccactc tctatgaaaa gtcctattaa 660
 atgacgtcat ttattttatt attgccggtt ggctgctcca cagccgcaat ttaattggatg 720
 gctgacacgg cacgaaaccg acgggcgggtg ccgtgggaat aattctagag taaaccctaac 780
 ggcgcggtta actttgacgg tggcgaagac gcgtggggat aggtggttgg tccgcgtgac 840
 ggcgcggtt cagcccgctg accttgagcc gagactataa atcgaggcga agggatgagc 900
 tttgccattg cgttcttctt ctgttcatct ctgaaattcg ggcggaatcc ttcttcttct 960
 caag 964

15
 <210> 13
 <211> 535
 <212> ADN
 <213> Lemna minor

<400> 13

```
gtatgcgtct ttcctccttg tgattcgatc tttctggttg ctagatctgg tctattgac 60
tgctctattg atctggctca tttatcgctg catcgggatc tattgatccg tatgttgatt 120
tgggatccgt aggttggttt ggatcggaga ctgcgatttg attcttgtga tttcgcttgg 180
atttcggaaa tcggtgtggt tgaagtcgtg cgatctttta gatctgctcc tttttttatt 240
tgctatttta tatttacgtt gtttatgatc gcggtattatt ttgattcggt tattcgagat 300
ccatgccgtt taactcgttc tttgtgctcc gatctttgcg atacgtcggg cgttctagat 360
ccgttcacta ggtagttttt aagtctcttg agcttgattt atatggattt gctgttttcc 420
aggaaaaatt tatgcgcgat tcttacgccg gtttcccat tttacttttag gtcgtgaatt 480
cttttgatct gagaatgatg aatctgacat gtaccttccg gtttgtaatt tgcag 535
```

5

<210> 14

<211> 980

<212> ADN

<213> Spirodela polyrrhiza

10

<400> 14

```
gtaaggagca gatctctttg atcggttttg ttcttctttt gttttgtttt tttttcttgc 60
ggatcttcgg ttgcatcatg ccttggtctg ttttattagt ttaggatatc ctcgtttgga 120

tctgagccga tcatatatgt taaagggttg gttcgatctc tttgttcatt ttgcgatgaa 180
aaggatgtat ccttttgatg tgaggcgatc ttctatggtt aagactttgt tcggtctatt 240
gatcatttct gttcttcgtt tttgagtttt tttctgcgga tatcgcatca tccctaggtt 300
tttgcttttg ttaggatgca tcctttggat ttgagccgat ctcccttggt taaggctgtg 360
tctgttgcag aggagaaagt ctgtcgaggt ccttatgcag gctttgtcca gatgcgcgtg 420
ctctctcatg ctatgaattt atgttttgag aactcctccc gggttttcta gatccggatt 480
tgaagtattc attgcggttc cccttcgggt ttatgtattt ctcgagttga tttggtccat 540
gatcgtgttc tgtccagatc tctcttgata tggatgagat attcgttacc tctttcaaac 600
atcggtggat gttcttttta gtcttggttc acctttatct agaaattaat tttcggtttg 660
aaacccctgc ttgttaagggt gatgtattcc ttctttatag atttcggtgt gttattttctt 720
aacggtgate tgtccgatcc atgtgttgca cctcttggtt tctgtgtaat cctctgtgaa 780
ttataattat gttttgaaaa cgtacttaag taaggggcat gttcccggtt taaaactttt 840
gttctatcaa tttgtggtta atagatcctg atttgtggtc gccttattct gtctttaatc 900
gtggatttta tttatcttga gcgcgtcctt ttcttttaaa atcatgtgtt taacctttca 960
gtcgtcatat gttccatcag 980
```

<210> 15

15

<211> 1104

<212> ADN

<213> Lemna aequinoctialis

<400> 15

20

```

gtatgggect cgatctttct gtttcaatcg agttttgac ttcgttttgg cggcgatcgg 60
tgttttcttt gtattgtgaa taaatccttg ataagaaaac cctagggttt gtgacctgtt 120
gacggatgcg tgcggatctg ttatttgtct tttaggcgat tttctcttgt ttgtaatagt 180
ttatcataac cagatgaaca tggatcaagt cgatttgact tattttttct gtgaaattag 240
gccgaaatcc ttttttttgg tttgagcctt gatatttcta tataattcga tttgattttt 300
tgttttcttc tgcgtctgat gctttctctt gactcctgat taaatttttg ctacggaaac 360
cctagatgtc gagatctgtt gacagattct ggcaaactcg tttttatcat aatcagatga 420
acgcaaatta agtcgatttg gtttttctct gaaattaggg gggaaactcc ttatagtatg 480
agcctcgata tttctataat agtcgatttg attttctctt gcctcctgat tcaatttttg 540
gtgcggaaac cctagatatt gtaatctgtt tacggatgct tgcggatctg atttttaata 600
ttgtgatcta ttgacggatg ctctgtagatc tggttgtttt gatttcttca tgccttatac 660
ggcgatttga ttcggcgatt aaaaattttc aattctttta aaaaaatat taagattttc 720
aacgtttcaa attatttcat agatcggcac aaatactttt catcagattc ctctgatgt 780
gatggtttgt gtttaaaatc tgttgaagat atcagattct attaggtcac cgatataatc 840
ttctctgttt attctgcgat cggtgcttac aaacctatt tcctacggtg attaattatt 900
tttaattccc tagctagcgt aatatatat ttttttaatt tgatctttgc attagtttcc 960
tccttttatt tgctattaat tgaaccgat gctacaaaac atcagatttt ttttcccaat 1020
tcgttgtcat cattatagaa aacttttata tgatattttt aatcgtcatt aatataattt 1080
tcaatttatt attttccctt gcag 1104

```

<210> 16

<211> 64

5 <212> ADN

<213> Lemna gibba

<400> 16

```

aagcacgagc tgagcgagaa ttcggggagg ctgagtcgaa gaggaagaga gaagtaggta 60
cgcc 64

```

10

<210> 17

<211> 58

<212> ADN

15 <213> Lemna gibba

<400> 17

```

actcgcaagt ggagagagga tccgagcgtc cagttagagg aagagagagg gaggcgcg 58

```

<210> 18

20 <211> 62

<212> ADN

<213> Lemna gibba

<400> 18

```

aaactccga ggtgagcaag gatccggagt cgagcgcgaa gaagagaaag agggaaagcg 60
cg 62

```

25

<210> 19
 <211> 1282
 <212> ADN
 <213> Lemma minor

5

<220>
 <221> CDS
 <222> (1)...(1276)

10

<400> 19

tgc gaa gcc tac ttc ggc aac agc ttc aac cgc cgg acg gag atg ctg	48
Cys Glu Ala Tyr Phe Gly Asn Ser Phe Asn Arg Arg Thr Glu Met Leu	
1 5 10 15	
aag aag gta gag gga aga gga tgg ttc cag tgc ctg tac agc gat act	96
Lys Lys Val Glu Gly Arg Gly Trp Phe Gln Cys Leu Tyr Ser Asp Thr	
20 25 30	
ctt cga agt tct gtt tgc cag gga ggg aat ttg cgg atg gac ccg gaa	144
Leu Arg Ser Ser Val Cys Gln Gly Gly Asn Leu Arg Met Asp Pro Glu	
35 40 45	
agg att agg atg tgc aaa ggg ggg gaa gat cta gag gag gtg atg aag	192
Arg Ile Arg Met Ser Lys Gly Gly Glu Asp Leu Glu Glu Val Met Lys	
50 55 60	
aga gag gag gaa gaa gaa ttg ccc aaa ttc gag gag ggg tgc ttc cag	240
Arg Glu Glu Glu Glu Glu Leu Pro Lys Phe Glu Glu Gly Ser Phe Gln	
65 70 75 80	
att gaa tct ggt tat gga agc gga ggg gaa gtt gga gag aga att gcg	288
Ile Glu Ser Gly Tyr Gly Ser Gly Gly Glu Val Gly Glu Arg Ile Ala	
85 90 95	
act gac gag gtc ctc gat aat gtt gtg cgc aaa ggc gct gtt cat gta	336
Thr Asp Glu Val Leu Asp Asn Val Val Pro Lys Gly Ala Val His Val	
100 105 110	
cat acc atg cgc aat ctc atc agt tgc att cag att gtt ggt ccc ggg	384
His Thr Met Arg Asn Leu Ile Ser Ser Ile Gln Ile Val Gly Pro Gly	
115 120 125	
cat ctt caa tgc tct cag tgg atc gac gaa ccg gtt ctt ctt gtc aca	432
His Leu Gln Cys Ser Gln Trp Ile Asp Glu Pro Val Leu Leu Val Thr	
130 135 140	
cgc ttc gaa tac gcc aat ctc ttt cac acc gtc acc gac tgg tac agc	480
Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser	
145 150 155 160	
gcc tac gca agc tgc agg att gcc aac ttg ccc tct cgc cct cac ttg	528
Ala Tyr Ala Ser Ser Arg Ile Ala Asn Leu Pro Ser Arg Pro His Leu	
165 170 175	
att ttc gtc gat ggc cat tgc agg gca gaa cag tta gag gac acg tgg	576
Ile Phe Val Asp Gly His Cys Arg Ala Glu Gln Leu Glu Asp Thr Trp	
180 185 190	
cga gcc ctg ttc tca acc gtc cga tac gcc aag aac ttc tcc cag cca	624
Arg Ala Leu Phe Ser Thr Val Arg Tyr Ala Lys Asn Phe Ser Gln Pro	
195 200 205	

```

gtc tgc ttc cgc cac gcc gtc ctc tcc cct ctt ggc tat gag aca gct 672
Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala
210 215 220

ctc ttc aaa ggc cta tca gag agc ttc agc tgt gag gga gtg ccg gcc 720
Leu Phe Lys Gly Leu Ser Glu Ser Phe Ser Cys Glu Gly Val Pro Ala
225 230 235 240

aat cag ctc aaa gtc aac cct gat gac cag aag act gcg aga ctg gct 768
Asn Gln Leu Lys Val Asn Pro Asp Asp Gln Lys Thr Ala Arg Leu Ala
245 250 255

gaa ttc gga gag atg atc agg gct gcc ttt gac ttt cct gtc gtt gac 816
Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Asp Phe Pro Val Val Asp
260 265 270

ccg ccc gtt gac ccg ttg acc aaa tcc atc ctc ttt gtg cgg cgg gaa 864
Pro Pro Val Asp Pro Leu Thr Lys Ser Ile Leu Phe Val Arg Arg Glu
275 280 285

gat tac gtg gcg cac cca cgc cac agt ggg aga gtg gag tcg cgg ttg 912
Asp Tyr Val Ala His Pro Arg His Ser Gly Arg Val Glu Ser Arg Leu
290 295 300

acc aat gag caa gag gtg ttt gac ttt ctg cac aaa tgg gca agt caa 960
Thr Asn Glu Gln Glu Val Phe Asp Phe Leu His Lys Trp Ala Ser Gln
305 310 315 320

cac aga agc agg tgc aac gtc agt gtg gtc aac ggg ctt ttc gcg cac 1008
His Arg Ser Arg Cys Asn Val Ser Val Val Asn Gly Leu Phe Ala His
325 330 335

atg gga atg aag gaa cag gtg aag gca att atg gaa gct tcg gtg gtg 1056
Met Gly Met Lys Glu Gln Val Lys Ala Ile Met Glu Ala Ser Val Val
340 345 350

gtc ggg gcc cac ggg gct ggt ttg act cat ctg gtg gca gca agg tca 1104
Val Gly Ala His Gly Ala Gly Leu Thr His Leu Val Ala Ala Arg Ser
355 360 365

acg aca gtt gtt ctt gag att ctg agc agt caa tat cgt aga ccg cac 1152
Thr Thr Val Val Leu Glu Ile Leu Ser Ser Gln Tyr Arg Arg Pro His
370 375 380

ttt caa ctg att tca cgg tgg aaa ggg ttg gac tac cac gca att aat 1200
Phe Gln Leu Ile Ser Arg Trp Lys Gly Leu Asp Tyr His Ala Ile Asn
385 390 395 400

ctt gcc ggg tcg tat gct gat cct cgg gag gtg gtc gag aaa ttg act 1248
Leu Ala Gly Ser Tyr Ala Asp Pro Arg Glu Val Val Glu Lys Leu Thr
405 410 415

ggc ata gtc gat ggg ctt gga tgt tga a gataag 1282
Gly Ile Val Asp Gly Leu Gly Cys *
420

```

<210> 20

<211> 1275

5 <212> ADN

<213> Lemna minor

<220>

<221> CDS

10 <222> (1)...(1276)

<221> característica_misc

<222> (0) ... (0)

<223> Xy1T isoforma #2; Codifica longitud parcial de beta-1, 2-xilosiltransferasa

<400> 20

5

tgc gaa gcc tac ttc ggc aac agc ttc aac cgc cgg acg gag atg ctg	48
Cys Glu Ala Tyr Phe Gly Asn Ser Phe Asn Arg Arg Thr Glu Met Leu	
1 5 10 15	
aag aag gta gag gga aga gga tgg ttc cag tgc ctg tac agc gat act	96
Lys Lys Val Glu Gly Arg Gly Trp Phe Gln Cys Leu Tyr Ser Asp Thr	
20 25 30	
ctt cga agt tct gtt tgc cag gga ggg aat ttg cgg atg gac ccg gaa	144
Leu Arg Ser Ser Val Cys Gln Gly Gly Asn Leu Arg Met Asp Pro Glu	
35 40 45	
agg att agg atg tcg aaa ggg ggg gaa gat cta gag gag gtg atg aag	192
Arg Ile Arg Met Ser Lys Gly Gly Glu Asp Leu Glu Glu Val Met Lys	
50 55 60	
aga gag gag gaa gaa gaa ttg ccc aaa ttc gag gag ggg tcg ttc cag	240
Arg Glu Glu Glu Glu Glu Leu Pro Lys Phe Glu Glu Gly Ser Phe Gln	
65 70 75 80	
att gaa tct ggt tat gga agc gga ggg gaa gtt gga gag aga att gcg	288
Ile Glu Ser Gly Tyr Gly Ser Gly Gly Glu Val Gly Glu Arg Ile Ala	
85 90 95	
act gac gag gtc ctc gat aat gtt gtg cgc aaa ggc gct gtt cat gta	336
Thr Asp Glu Val Leu Asp Asn Val Val Pro Lys Gly Ala Val His Val	
100 105 110	
cat acc atg cgc aat ctc atc agt tcg att cag att gtt ggt ccc ggg	384
His Thr Met Arg Asn Leu Ile Ser Ser Ile Gln Ile Val Gly Pro Gly	
115 120 125	
cat ctt caa tgc tct cag tgg atc gac gaa cgc gtt ctt ctt gtc aca	432
His Leu Gln Cys Ser Gln Trp Ile Asp Glu Pro Val Leu Leu Val Thr	
130 135 140	
cgc ttc gaa tac gcc aat ctc ttt cac acc gtc acc gac tgg tac agc	480
Arg Phe Glu Tyr Ala Asn Leu Phe His Thr Val Thr Asp Trp Tyr Ser	
145 150 155 160	
gcc tac gca agc tcg agg att gcc aac ttg ccc tct cgc cct cac ttg	528
Ala Tyr Ala Ser Ser Arg Ile Ala Asn Leu Pro Ser Arg Pro His Leu	
165 170 175	
att ttc gtc gat ggc cat tgc agg gca gaa cag tta gag gac acg tgg	576
Ile Phe Val Asp Gly His Cys Arg Ala Glu Gln Leu Glu Asp Thr Trp	
180 185 190	
cga gcc ctg ttc tca acc gtc cga tac gcc aag aac ttc tcc cag cca	624
Arg Ala Leu Phe Ser Thr Val Arg Tyr Ala Lys Asn Phe Ser Gln Pro	
195 200 205	
gtc tgc ttc cgc cac gcc gtc ctc tcc cct ctt ggc tat gag aca gct	672
Val Cys Phe Arg His Ala Val Leu Ser Pro Leu Gly Tyr Glu Thr Ala	
210 215 220	
ctc ttc aaa ggc cta tca gag agc ttc agc tgt gag gga gtg ccg gcc	720
Leu Phe Lys Gly Leu Ser Glu Ser Phe Ser Cys Glu Gly Val Pro Ala	

ES 2 395 842 T3

225	230	235	240	
aat cag ctc aaa gtc aac cct gat gac cag aag act gcg aga ctg gct	Asn Gln Leu Lys Val Asn Pro Asp Asp Gln Lys Thr Ala Arg Leu Ala	768		
	245	250	255	
gaa ttc gga gag atg atc agg gct gcc ttt gac ttt cct gtc gtt gac	Glu Phe Gly Glu Met Ile Arg Ala Ala Phe Asp Phe Pro Val Val Asp	816		
	260	265	270	
ccg ccc gtt gac ccg ttg acc aaa tcc atc ctc ttt gtg cgg cgg gaa	Pro Pro Val Asp Pro Leu Thr Lys Ser Ile Leu Phe Val Arg Arg Glu	864		
	275	280	285	
gat tac gtg gcg cac cca cgc cac agt ggg aga gtg gag tcg cgg ttg	Asp Tyr Val Ala His Pro Arg His Ser Gly Arg Val Glu Ser Arg Leu	912		
	290	295	300	
acc aat gag caa gag gtg ttt gac ttt ctg cac aaa tgg gca agt caa	Thr Asn Glu Gln Glu Val Phe Asp Phe Leu His Lys Trp Ala Ser Gln	960		
	305	310	315	320
cac aga agc agg tgc aac gtc agt gtg gtc aac ggg ctt ttc gcg cac	His Arg Ser Arg Cys Asn Val Ser Val Val Asn Gly Leu Phe Ala His	1008		
	325	330	335	
atg gga atg aag gaa cag gtg aag gca att atg gaa gct tcg gtg gtg	Met Gly Met Lys Glu Gln Val Lys Ala Ile Met Glu Ala Ser Val Val	1056		
	340	345	350	
gtc ggg gcc cac ggg gct ggt ttg act cat ctg gtg gca gca agg tca	Val Gly Ala His Gly Ala Gly Leu Thr His Leu Val Ala Ala Arg Ser	1104		
	355	360	365	
acg aca gtt gtt ctt gag att ctg agc agt caa tat cgt aga ccg cac	Thr Thr Val Val Leu Glu Ile Leu Ser Ser Gln Tyr Arg Arg Pro His	1152		
	370	375	380	
ttt caa ctg att tca cgg tgg aaa ggg ttg gac tac cac gca att aat	Phe Gln Leu Ile Ser Arg Trp Lys Gly Leu Asp Tyr His Ala Ile Asn	1200		
	385	390	395	400
ctt gcc ggg tcg tat gct gat cct cgg gag gtg gtc gag aaa ttg act	Leu Ala Gly Ser Tyr Ala Asp Pro Arg Glu Val Val Glu Lys Leu Thr	1248		
	405	410	415	
ggc ata gtc gat ggg ctt gga tgt tga	Gly Ile Val Asp Gly Leu Gly Cys *	1275		
	420			

<210> 21

<211> 424

5 <212> PRT

<213> Lemna minor

<400> 21

Cys	Glu	Ala	Tyr	Phe	Gly	Asn	Ser	Phe	Asn	Arg	Arg	Thr	Glu	Met	Leu
1				5					10					15	
Lys	Lys	Val	Glu	Gly	Arg	Gly	Trp	Phe	Gln	Cys	Leu	Tyr	Ser	Asp	Thr
		20						25					30		
Leu	Arg	Ser	Ser	Val	Cys	Gln	Gly	Gly	Asn	Leu	Arg	Met	Asp	Pro	Glu
		35					40					45			
Arg	Ile	Arg	Met	Ser	Lys	Gly	Gly	Glu	Asp	Leu	Glu	Glu	Val	Met	Lys
	50					55					60				
Arg	Glu	Glu	Glu	Glu	Glu	Leu	Pro	Lys	Phe	Glu	Glu	Gly	Ser	Phe	Gln
65					70					75					80
Ile	Glu	Ser	Gly	Tyr	Gly	Ser	Gly	Gly	Glu	Val	Gly	Glu	Arg	Ile	Ala
			85						90					95	
Thr	Asp	Glu	Val	Leu	Asp	Asn	Val	Val	Pro	Lys	Gly	Ala	Val	His	Val
			100					105						110	
His	Thr	Met	Arg	Asn	Leu	Ile	Ser	Ser	Ile	Gln	Ile	Val	Gly	Pro	Gly
		115					120					125			
His	Leu	Gln	Cys	Ser	Gln	Trp	Ile	Asp	Glu	Pro	Val	Leu	Leu	Val	Thr
	130					135					140				
Arg	Phe	Glu	Tyr	Ala	Asn	Leu	Phe	His	Thr	Val	Thr	Asp	Trp	Tyr	Ser
145					150					155					160
Ala	Tyr	Ala	Ser	Ser	Arg	Ile	Ala	Asn	Leu	Pro	Ser	Arg	Pro	His	Leu
				165					170					175	
Ile	Phe	Val	Asp	Gly	His	Cys	Arg	Ala	Glu	Gln	Leu	Glu	Asp	Thr	Trp
			180					185						190	
Arg	Ala	Leu	Phe	Ser	Thr	Val	Arg	Tyr	Ala	Lys	Asn	Phe	Ser	Gln	Pro
		195					200					205			
Val	Cys	Phe	Arg	His	Ala	Val	Leu	Ser	Pro	Leu	Gly	Tyr	Glu	Thr	Ala
	210					215					220				
Leu	Phe	Lys	Gly	Leu	Ser	Glu	Ser	Phe	Ser	Cys	Glu	Gly	Val	Pro	Ala
225					230					235					240
Asn	Gln	Leu	Lys	Val	Asn	Pro	Asp	Asp	Gln	Lys	Thr	Ala	Arg	Leu	Ala
				245					250					255	
Glu	Phe	Gly	Glu	Met	Ile	Arg	Ala	Ala	Phe	Asp	Phe	Pro	Val	Val	Asp
			260					265					270		
Pro	Pro	Val	Asp	Pro	Leu	Thr	Lys	Ser	Ile	Leu	Phe	Val	Arg	Arg	Glu
		275					280					285			
Asp	Tyr	Val	Ala	His	Pro	Arg	His	Ser	Gly	Arg	Val	Glu	Ser	Arg	Leu
	290					295					300				
Thr	Asn	Glu	Gln	Glu	Val	Phe	Asp	Phe	Leu	His	Lys	Trp	Ala	Ser	Gln
305					310					315					320
His	Arg	Ser	Arg	Cys	Asn	Val	Ser	Val	Val	Asn	Gly	Leu	Phe	Ala	His
				325					330					335	
Met	Gly	Met	Lys	Glu	Gln	Val	Lys	Ala	Ile	Met	Glu	Ala	Ser	Val	Val
			340					345					350		
Val	Gly	Ala	His	Gly	Ala	Gly	Leu	Thr	His	Leu	Val	Ala	Ala	Arg	Ser
		355					360					365			
Thr	Thr	Val	Val	Leu	Glu	Ile	Leu	Ser	Ser	Gln	Tyr	Arg	Arg	Pro	His
	370					375						380			
Phe	Gln	Leu	Ile	Ser	Arg	Trp	Lys	Gly	Leu	Asp	Tyr	His	Ala	Ile	Asn
385					390					395					400
Leu	Ala	Gly	Ser	Tyr	Ala	Asp	Pro	Arg	Glu	Val	Val	Glu	Lys	Leu	Thr
				405					410					415	
Gly	Ile	Val	Asp	Gly	Leu	Gly	Cys								
			420												

<210> 22

<211> 28

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> cebador

<400> 22

5 atggctgact gctgctggcg ctctcaac 28

<210> 23

<211> 28

<212> ADN

10 <213> Secuencia Artificial

<220>

<223> cebador

15 <400> 23

atgtctagaa tgcagcagca agtcacc 28

<210> 24

<211> 33

20 <212> ADN

<213> Secuencia Artificial

<220>

<223> cebador

25

<400> 24

atgactagtt gccaagccta ctccggcaac agc 33

<210> 25

30 <211> 30

<212> ADN

<213> Secuencia Artificial

<220>

35 <223> cebador

<400> 25

atgggatcgg aatctcaaga acaactgtcg 30

40 <210> 26

<211> 33

<212> ADN
 <213> Secuencia Artificial

 <220>
 5 <223> cebador

 <400> 26
 atgggtacct gcgaagccta ctctggcaac agc 33

 <210> 27
 10 <211> 29
 <212> ADN
 <213> Secuencia Artificial

 <220>
 15 <223> cebador

 <400> 27
 atgggatcca ctggctggga gaagttctt 29

 20 <210> 28
 <211> 28
 <212> ADN
 <213> Secuencia Artificial

 25 <220>
 <223> cebador

 <400> 28
 atggagctct gctgctggg ctctcaac 28
 30
 <210> 29
 <211> 28
 <212> ADN
 <213> Secuencia Artificial
 35
 <220>
 <223> cebador

 <400> 29
 40 atgggtacca tgcagcagca agtgcacc 28

REIVINDICACIONES

- 1.- Una planta de lenteja de agua o una célula o nódulo de lenteja de agua que expresa una glicoproteína recombinante que tiene un perfil de *N*-glicosilación sustancialmente homogéneo, en donde al menos el 90% de las especies de *N*-glicano presentes en el perfil son glicanos G0, y en donde dicha planta, célula vegetal o nódulo comprende al menos una construcción de nucleótido recombinante que proporciona la inhibición de la expresión de α 1,3-fucosiltransferasa (FucT) y β 1,2-xilosiltransferasa (XylT) en dicha planta, célula vegetal o nódulo por medio de:
- (a) la supresión o co-supresión sentido a través de una casete de expresión que expresa una molécula de ARN correspondiente a parte o todo el ARN mensajero que codifica una FucT o XylT, o ambas, en la orientación sentido;
 - (b) la supresión antisentido a través de una casete de expresión que expresa una molécula de ARN complementaria a parte o a todo el ARN mensajero que codifica una FucT o XylT, o ambas, en la orientación antisentido;
 - (c) la interferencia de ARN de cadena doble, en donde el ARN sentido de (a) y el ARN antisentido de (b) son expresados en la misma célula;
 - (d)
 - (i) la interferencia de ARN de horquilla (ARNhp), o
 - (ii) la interferencia de ARN de horquilla que contiene intrón (ARNihp)
 a través de una casete de expresión que expresa una molécula de ARN que se hibrida consigo misma para formar una estructura de horquilla que comprende una región de lazo de cadena sencilla y un tallo de bases emparejadas, en donde la región de tallo de bases emparejadas comprende una secuencia sentido correspondiente a parte o a toda la secuencia diana de FucT o XylT, y una secuencia antisentido que es parcial o totalmente complementaria a la secuencia sentido; o
 - (e) la inhibición de microARN (miARN) a través de una construcción que expresa un ARN que forma una estructura de horquilla que contiene una secuencia de 22 nucleótidos que es complementaria a la secuencia diana de FucT o XylT y que contiene 22 nucleótidos de dicha secuencia de FucT o XylT en orientación sentido y 21 nucleótidos de una secuencia antisentido correspondiente que es complementaria a la secuencia sentido.
- 2.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de la reivindicación 1, en donde dichos *N*-glicanos están desprovistos de dichos residuos de fucosa y xilosa.
- 3.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de la reivindicación 1 o de la reivindicación 2, en donde dicha glicoproteína es una proteína de mamífero de interés.
- 4.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de una cualquiera de las reivindicaciones 1 a 3, en donde dicha glicoproteína es un anticuerpo monoclonal de interés.
- 5.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de la reivindicación 4, en donde dicho anticuerpo monoclonal tiene una afinidad de unión para un FcyRIII incrementada, una actividad de citotoxicidad celular dependiente de anticuerpo (ADCC) incrementada, una actividad de citotoxicidad dependiente de complemento (CDC) disminuida, o cualquier combinación de las mismas, como resultado de dicho modelo de *N*-glicosilación alterado de dicho anticuerpo monoclonal.
- 6.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua según una cualquiera de las reivindicaciones 1 a 5, en donde dicha planta, célula vegetal o nódulo comprende una construcción de nucleótidos integrada establemente en su genoma, en donde la construcción se selecciona del grupo que consiste en:
- (a) una construcción de nucleótido que comprende una primera secuencia de nucleótidos que es capaz de inhibir la expresión o la función de una α 1,3-fucosiltransferasa (FucT) en una planta, y una segunda secuencia de nucleótidos que es capaz de inhibir la expresión o la función de una β 1,2-xilosiltransferasa (XylT) en una planta, en donde dicha primera secuencia de nucleótidos y dicha segunda secuencia de nucleótidos están ligadas operativamente a al menos un promotor que es funcional en una célula vegetal; y
 - (b) una construcción de nucleótidos que comprende un polinucleótido de fusión que es capaz de inhibir la expresión o la función de una α 1,3-fucosiltransferasa (FucT) y una β 1,2-xilosiltransferasa (XylT) en una planta, en donde dicho polinucleótido de fusión está ligado operativamente a un promotor que es funcional en una célula vegetal.
- 7.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua según una cualquiera de las reivindicaciones 1 a 5, en donde dicha planta, célula vegetal o nódulo comprende:

(a) una primera construcción de nucleótidos que comprende una secuencia de polinucleótido que es capaz de inhibir la expresión o la función α 1,3-fucosiltransferasa (FucT) en una planta, en donde dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha FucT comprende en la orientación 5'-a-3' y ligados operativamente:

- 5 i. un fragmento de FucT directo, comprendiendo dicho fragmento de FucT directo entre aproximadamente 500 y aproximadamente 800 nucleótidos contiguos que tienen al menos un 90% de identidad de secuencia con respecto a una secuencia de nucleótidos de aproximadamente 500 a aproximadamente 800 nucleótidos contiguos de un polinucleótido que codifica dicha FucT;
- 10 ii. una secuencia espaciadora que comprende de aproximadamente 200 a aproximadamente 700 nucleótidos; y
- 15 iii. un fragmento de FucT inverso, presentando dicho fragmento de FucT inverso una longitud suficiente y una complementariedad suficiente respecto a dicho fragmento de FucT directo como para que dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha FucT sea transcrita como una molécula de ARN capaz de formar una estructura de ARN de horquilla; en donde dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha FucT está ligada operativamente a un promotor que es funcional en una célula vegetal; y

(b) una segunda construcción de nucleótidos que comprende una secuencia de polinucleótido que es capaz de inhibir la expresión o la función β 1,2-xilosiltransferasa (XylT) en una planta, en donde dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha XylT comprende en la orientación 5'-a-3' y ligados operativamente:

- 20 i. un fragmento de XylT directo, comprendiendo dicho fragmento de XylT directo entre aproximadamente 500 y aproximadamente 800 nucleótidos contiguos que tienen al menos un 90% de identidad de secuencia con respecto a una secuencia de nucleótidos de aproximadamente 500 a aproximadamente 800 nucleótidos contiguos de un polinucleótido que codifica dicha XylT;
- 25 ii. una secuencia espaciadora que comprende de aproximadamente 200 a aproximadamente 700 nucleótidos; y
- 30 iii. un fragmento de XylT inverso, presentando dicho fragmento de XylT inverso una longitud suficiente y una complementariedad suficiente respecto a dicho fragmento de XylT directo como para que dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha XylT sea transcrita como una molécula de ARN capaz de formar una estructura de ARN de horquilla; en donde dicha secuencia de polinucleótido que es capaz de inhibir la expresión o la función de dicha XylT está ligada operativamente a un promotor que es funcional en una célula vegetal.

8.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de una cualquiera de las reivindicaciones 1 a 7, en donde dicha lenteja de agua pertenece a un género seleccionado del grupo que consiste en el género *Spirodela*, género *Wolffia*, género *Wolffiella*, género *Landoltia* y género *Lemna*.

9.- La planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de la reivindicación 8, en donde dicha lenteja de agua es un miembro de una especie seleccionada del grupo que consiste en *Lemna minor*, *Lemna miniscula*, *Lemna aequinoctialis* y *Lemna gibba*.

10.-Un método para usar la planta de lenteja de agua o la célula vegetal o nódulo de lenteja de agua de una cualquiera de las reivindicaciones 1 a 9 para producir una composición de glicoproteínas que tiene un perfil de *N*-glicosilación sustancialmente homogéneo, en donde al menos el 90% de las especies de *N*-glicano presentes en el perfil son glicanos G0, comprendiendo dicho método:

- (a) cultivar dicha planta de lenteja de agua o célula vegetal o nódulo de lenteja de agua en un medio de cultivo en las condiciones que son adecuadas para la expresión de dicha glicoproteína; y,
- 45 (b) recolectar dicha glicoproteína de dicha planta de lenteja de agua, célula vegetal o nódulo, o dicho medio de cultivo;

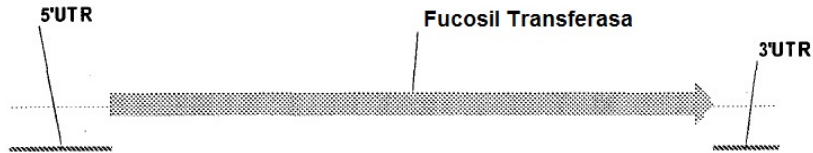
produciendo con ello dicha composición de glicoproteína.

11.-El método de la reivindicación 10, en el que dicha glicoproteína es un anticuerpo monoclonal.

12.-El método de la reivindicación 10 u 11, en donde dicho anticuerpo monoclonal tiene una afinidad de unión a un FcyRIII incrementada, una citotoxicidad celular dependiente de anticuerpo (ADCC) incrementada, una actividad citotoxicidad dependiente de complemento (CDC) disminuida, o cualquier combinación de las mismas, como resultado de dicho modelo de *N*-glicosilación alterado de dicho anticuerpo monoclonal.

Información de ADN y secuencia de aminoácidos

alfa 1-3 fucosiltransferasa de *Lemna minor*:



RACE PCR de fucosiltransferasa de *L. minor*
1865 pb

Secuencia de ADNc de isoforma nº1 de FucT: (UTRs en fuente estándar, región codificadora en negrita)

acgcggggggaagtgggttagtagctcagtggaatgtctattagagggggaagaggggagggatccgagggg
aacgaggaaggtgtgcccgaattctcgtagatttcttcaattcctgcagatctcgtctctctctgatttcttcccagagctc
cgcccgtaggaactcaatcggaactcgatccaagttgacgagggcctacygaggaaggcgattttccgaagcctgcgacga
tgccacactctgctgctggtgctctcaacgcgggtggcagggtcgggggcaggagaggtgggtcagattgcttcccttct
ttgtgttgatgctggtggttaggggagatctggttctcggggcggtggatgtggtcaagaacgcgctatggttcaaaact
ggacttcttccacttgtttttcttaccagtttcttctacacgtggtccgagaccgtcaaggaggaagaggattgcaagg
actggctggaagagtagatgctggtcgattacaagagagatttccggtggaacccgttctggttaaatagcgctgaacagg
attggagttcatgttcagtggtggtgtaagttcggtatcattcccggaagaacgcctgtagctacatttgggttctctcaga
atccatcaacagtcagtgctcatcgatccatggaatcatccattatttattggagaataatcttgataatgcacgacgga
aaggctatcaaatgtgtagacaactagctcttctgctcagatgtgctgctggttatttctcatgggctgaatatgatata
tgccgctcttccagccgaaaactgctggtgcacttgcgtgctgcatattatcttaattgaggagcacgtaatttccgcttgc
aggcccttgatagctcgaagtcgaatattaagattgattcatatggtgcttgcctatcgcaaccaagacggttaaagtgg
acaaggtacaaactttgaagcgggtataagttcagcttagcttttgaaaactcgaacgaggatgactatgttactgagaagt
tctttcaatctcttgcgtgagctattcctgttgcgtgagggccccaacattcaaaatttgcgcatctcttgatt
caattctgcacatcagggagcccaaggatgtcagttcagtcgctgagagaaatgaaatttctcgttcaaatccagaagcat
ataaccaatcactgaggtggaagtttgaggcccttctaactccttcaaagccctggtggacatggcagcagttcactcct
cctgcgcctatgcattcacattgccaccaagatcagagagagagagagaaacccgaatttcaagactcgccttgca
agtgcacccgcaatgggtctaccttatatcacttatacgcgcgcgaaagaggcaccttgacttcttataatcttcatga
gatcggataatctatcactgaaagcgtgggtcaacagttcttgagaaattcagttcttgaagcacgtgcccatttgga
agaaggagagggccagagagctgaaaggaggagcaagctggatctttacagaatctatccagtgggcattactcagagag
aagctctcttcttcttccagttcaacactgacaaagaacttcaaatctaccttgaatcccatccatgtgcaagtttgaag
tcatctttatttgatccctgaggttaattaggtcacgaattcagctaatttggttaattatgcttcaagccacatggtatt
tcatatcattaattgaaggcatagtttagttgatattgacattttcgtctaggatcattctaaagctatcccaatgaactt
aa

Secuencia de aminoácidos predicha para la isoforma nº1 de FucT:

matsaagalnaggrvgrsrsvrllpffvmlvvgeiwlgrldvknmaamvqnwtsshlfflpvssytwtsetvkeeedck
dwlervdavdykrdfrvepvlvndaeqdwsscsvgckfsgfgrtpdatfgfsqnpstvsvhsmesshyllennldnarr
kgyqivmttllsdvpvgyswaeydimaplpktagalvaafisncgarnfrlqaldmleksnikidsyagachrnqdgkv
dkvqtlkrykfslafensneddyvtekffqslvagaaipvvvafisncgarnfripssdsilhirepkdvssvaermkflasnpa
ynqslrwkfegpsnsfkalvmaavhsscrclhiatkirekeernpnfkrpckctrngstlyhlyarergtfdflsifm
rsdnslslkalgstvlekfsslkhpwiwkerpeslkggskldlyriypvgitqrealfsqfntdkelqiyleshpcake
vifi*

FIG. 1

Alineamiento de secuencias de aminoácidos de alfa 1,3-fucosiltransferasa procedente de *Lemna minor* y otras plantas superiores

Alineamiento total positivo -- 89,2%

Identidad de alineamiento total -- 41,1%

Triticum aestivum: N° de acceso GenBank AJ582182.1

Hordeum vulgare: N° de acceso GenBank AJ582181.1

Oryza sativa: N° de acceso GenBank XM_482907.1

Medicago sativa: N° de acceso GenBank AY082444.1

Arabidopsis thaliana: N° de acceso GenBank AY184990.1

	1	35
Lemna minor FucT	(1) -----MATSAGALNAGGRVG	
Triticum aestivum FucT	(1) -----MKGSHSOS-----QAAS	
Hordeum vulgare FucT	(1) -----MESHSLT-----QAAS	
Oryza sativa FucT	(1) -----MKGSHSOSONORQSCAG	
Medicago sativa FucT	(1) MGLVSRT-----TTTTTQEGLEFVSVVSTTV	
Arabidopsis thaliana FucT	(1) MGVFSNLRGPRAGATHDEFPATNGSPSSSSSPSS	
Consenso	(1) -----GSHS S S A AAS	
	36	70
Lemna minor FucT	(17) GRSWVRIIPFFVLNLVGEIWFELGRLDVKNAM	
Triticum aestivum FucT	(13) RRRRCGWLLPLLVGVAFVGEIAFLGRLDMAKNAAS	
Hordeum vulgare FucT	(12) RRRRCGWLLPLLVGVAFVGEIAFLGRLDMAKNAAS	
Oryza sativa FucT	(18) RRRRCGWLLPLLVGAFLAETFLGRLDMAKNAAS	
Medicago sativa FucT	(27) PKKWSNLMPLFVAIVVIAETFLGRLDMAKNAAM	
Arabidopsis thaliana FucT	(36) IKRLSNLPLCVAVVIAETFLGRLDKVALVDT	
Consenso	(36) RRRRCGWLLPLLVGLAVIAETFLGRLDMAKNAAS	
	71	105
Lemna minor FucT	(52) VQNTSSHLFFLVSS-----YTWSETVK	
Triticum aestivum FucT	(48) VESWTTSFYRLSSTWG-----ADAPP	
Hordeum vulgare FucT	(47) VESWTTSFHRLSSTWG-----ADAPP	
Oryza sativa FucT	(53) VESWTTSLYARSSAPPRDGKAAVVVPGADALDAPP	
Medicago sativa FucT	(62) VADLFYRSRAVVEGDD-----FGL	
Arabidopsis thaliana FucT	(71) ETDNFKQSPQLSQ-----APP	
Consenso	(71) VESWTTSFYALSS G ADAPP	
	106	140
Lemna minor FucT	(76) -----EELCKDNLIRVDAVDYKROEPVE	
Triticum aestivum FucT	(69) GSG-----DDEECEERLEPEDAVPYDRDFEKD	
Hordeum vulgare FucT	(68) GSG-----DDEECEERLEPEDAVPYDRDFEKD	
Oryza sativa FucT	(88) GCGEVVEEDDGIPLCEERLEPEDGVPHDPDFDKD	
Medicago sativa FucT	(81) EIVGGDKNLELEPETCEENLEPEDAVTYSRNTNKE	
Arabidopsis thaliana FucT	(87) RRSDRKKIGLFTDSCLEENLMEDSVTYSKDETKD	
Consenso	(106) GSG DDEECEEWLEREDAVPYDRDFEKD	

FIG. 2A

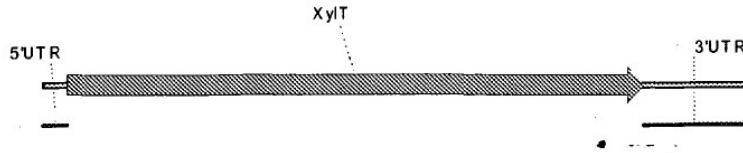
		141		175
Lemna minor FucT	(100)	PVLVGLAECDYSSCSVCEFGSFPGKTPDATFGFS		
Triticum aestivum FucT	(97)	PVLVGGAAKDWNRCVGCCEFGEPASKTPDATFGIA		
Hordeum vulgare FucT	(96)	PVLVGGAAKDWNRCVGCCEFGEPASKTPDATFGIA		
Oryza sativa FucT	(123)	PVLVGGAAKDWNRCVGCCEFGESATKTPDATFGIA		
Medicago sativa FucT	(116)	PMFYGGABQDAKSCSVCECKEGINGDRKFFAAFGLP		
Arabidopsis thaliana FucT	(122)	PMFYSCGEKDFQWCSVLCTFGDSSCKTPDAAFGLG		
Consenso	(141)	PVLVGGAEKDWNRCSVGCCEFGF ASKTPDATFGIA		
		176		210
Lemna minor FucT	(135)	QNPSTVSLHRSMESSHYYLENNIDNARRRGYQIVM		
Triticum aestivum FucT	(132)	QDPVESILRSMESSQYYSENNINAARGRGYQIVM		
Hordeum vulgare FucT	(131)	QDPVESILRSMESSQYYSENNINAARGRGYQIVM		
Oryza sativa FucT	(158)	QDPVESILRSMESSQYYSENNIAVARGRGYQIVM		
Medicago sativa FucT	(151)	QDAGTASVLRSMESAGYYAENMLAMARRGYHIVM		
Arabidopsis thaliana FucT	(157)	QDGTLSILRSMESSQYYPENDLAQARRGYDIVM		
Consenso	(176)	QDPVESILRSMESSQYYSENNIA ARRRGYQIVM		
		211		245
Lemna minor FucT	(170)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Triticum aestivum FucT	(167)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Hordeum vulgare FucT	(166)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Oryza sativa FucT	(193)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Medicago sativa FucT	(186)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Arabidopsis thaliana FucT	(192)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
Consenso	(211)	TTSLSLSDVPVGYFSWAEDYDIMPVPPKTEALAAA		
		246		280
Lemna minor FucT	(205)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Triticum aestivum FucT	(202)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Hordeum vulgare FucT	(201)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Oryza sativa FucT	(228)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Medicago sativa FucT	(221)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Arabidopsis thaliana FucT	(227)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
Consenso	(246)	FISNCGARNFRLQALEMLESLLVIDSYGGSCHRRN		
		281		315
Lemna minor FucT	(240)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Triticum aestivum FucT	(237)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Hordeum vulgare FucT	(236)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Oryza sativa FucT	(263)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Medicago sativa FucT	(256)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Arabidopsis thaliana FucT	(262)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
Consenso	(281)	DGKVDKVELTKRYKPSLAFENSNEEDYVTEKFFQS		
		316		350
Lemna minor FucT	(275)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Triticum aestivum FucT	(272)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Hordeum vulgare FucT	(271)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Oryza sativa FucT	(298)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Medicago sativa FucT	(291)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Arabidopsis thaliana FucT	(297)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		
Consenso	(316)	LVTGAIPVVVGAPNIQEFAPGEDAILHIKELDDV		

FIG. 2B

		351	385
Lemna minor FucT	(310)	SVAKRMKFDASNEEAFNOSLRWKYEGPSNDFKALV	
Triticum aestivum FucT	(307)	SVAKTMKHIAENEDAFNOSLRWKYEGPSDFKALV	
Hordeum vulgare FucT	(306)	SVAKTMKHIAENEDAFNOSLRWKYEGPSDFKALV	
Oryza sativa FucT	(333)	SVAKTMKHIAENEDAFNOSLRWKYEGPSDFKALV	
Medicago sativa FucT	(326)	SVAKSMRYIAENEDAFNHSLRWKYEGPSDFKALV	
Arabidopsis thaliana FucT	(332)	PVAKRMKFIAPNPAANOTLRWKYEGPSDFKALV	
Consenso	(351)	SVAKTMKHIAENPEAFNOSLRWKYEGPSDFKALV	
		386	420
Lemna minor FucT	(345)	DMAAVHSSCRLCIHIAATKIREKEERNENFKTRECK	
Triticum aestivum FucT	(342)	DMAAVHSSCRLCIHIAATKIREKEERTPKFMRSCS	
Hordeum vulgare FucT	(341)	DMAAVHSSCRLCIHIAATKIREKEERTPKFMRSCS	
Oryza sativa FucT	(368)	DMAAVHSSCRLCIHIAATKIREKEERTPKFMRSCS	
Medicago sativa FucT	(361)	DMAAVHSSCRLCIHIAATKIREKEERTPKFMRSCS	
Arabidopsis thaliana FucT	(367)	DMAAVHSSCRLCIFIAATKIREKEERTPKFMRSCS	
Consenso	(386)	DMAAVHSSCRLCIHIAATKIREKEERTPKFMRSCS	
		421	455
Lemna minor FucT	(380)	CERN-GSIVYHLYANERGTIDFLSIFNPSDNLSEK	
Triticum aestivum FucT	(377)	CSSK-RGTVYHLEVRERGRFKTESIFLRSDQLTIG	
Hordeum vulgare FucT	(376)	CSSK-RGTVYHLEVRERGRFKTESIFLRSDQLTIG	
Oryza sativa FucT	(403)	CSSK-RGTVYHLEVRERGRFKTESIFLRSDQLTIG	
Medicago sativa FucT	(396)	CVRG-SETVYHLYVREPGTFEMESIYLRSSNLTIE	
Arabidopsis thaliana FucT	(402)	CVRGSDIVYHLYVREPGTFEMESIYLRSSNLTIE	
Consenso	(421)	CSSK RGTVYHLEVRERGRFKTESIFLRSDQLTIG	
		456	490
Lemna minor FucT	(414)	ALESVLAHFKSLKHPVWVWDERPPSIRGGDELKV	
Triticum aestivum FucT	(411)	ALESVLAHFKSLKHPVWVWDERPPSIRGGDELKV	
Hordeum vulgare FucT	(410)	ALESVLAHFKSLKHPVWVWDERPPSIRGGDELKV	
Oryza sativa FucT	(437)	ALESVLAHFKSLKHPVWVWDERPPSIRGGDELKV	
Medicago sativa FucT	(430)	SEFIAVITKFSLSINNVVWVWDERPPSIRGGDELKV	
Arabidopsis thaliana FucT	(437)	ALESVLAHFKSLKHPVWVWDERPPSIRGGDELKV	
Consenso	(456)	ALESVLAHFKSLKHPVWVWDERP SIKGGDELKV	
		491	525
Lemna minor FucT	(449)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Triticum aestivum FucT	(446)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Hordeum vulgare FucT	(445)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Oryza sativa FucT	(472)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Medicago sativa FucT	(465)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Arabidopsis thaliana FucT	(472)	YRIYPIGLTORQALYQFQDDAELARYIKHPC	
Consenso	(491)	YRIYPIGLTORQALY QFQ DDAELARYIK HPC	
		526	
Lemna minor FucT	(484)	KFEVIFV-	
Triticum aestivum FucT	(481)	KFEVIFV-	
Hordeum vulgare FucT	(480)	KFEVIFV-	
Oryza sativa FucT	(507)	KFEVIFV-	
Medicago sativa FucT	(500)	KFEVIFV-	
Arabidopsis thaliana FucT	(507)	KFEVIFV-	
Consenso	(526)	KFEVIFV	

FIG. 2C

Beta 1-2 xilosiltransferasa de *Lemna minor*



beta 1-2 xilosiltransferasa de *Lemna minor*
1860 pb

Secuencia de ADN de isoforma nº1 de XylT: (UTRs en fuente estándar, región codificadora en negrita)

ggcttccaaccggaggatctcgagctgaagaatcttcagactgaagaattcatgtgatcccatggccttgggtgaactcgc
gagggagcagggtcagacgcacgcgaagcccaccttcgttttccctcttgatcaacgtagtctgtctcctgtacttttcc
gtcagaaccctaattccattcccgacgcttgtcttcacggggaatgcgacaaacccccgatttttagtgactccccggcgat
ggaacttgaagccatggcgattcttccctcccttctgccaatgggtgcccagctcccaccctgccagggtcctcgcgaag
cctacttcggcaacagcttcaaccgccggacggagatgctgaagaaggtagaggaagaggaatggttccagtgccctgtaca
gcgatactcttcgaagtctgtttgcccaggaggggaatttgcgagtgaccggaaaggattaggaatgcgaaggggggg
aagatctagaggaggtgatgaagagagaggaggaagaagaattgcccgaattcgaggaggggtcgttccagattgaatctg
gttatggaagcggagggggaagtggagagagaattgcgactgacgaggtcctcgataatgttgtgcccgaaggcgtgttc
atgtacataccatgcgcaatctcatcagttcgatttcagattgttgggtcccgggcatcttcaatgctctcagtggaatcgacg
aacgggttcttcttgcacacgcttcgaatacgcgaatctcttccacacgctcaccgactgggtacagccctacgcaagct
cgaggattgccaacttgcccttctgcgcctcacttaattttcgtcgatggccattgcagggcggaacagttagaggacatgt
ggagagccctgttctcgacgctccgatactccaagaacttctcccagccaatctgcttcccgcacgctcgtctctcacctc
tggtctatgagacggctctcttcaaggccctatcagagagcttcagctgtgagggagctccggccaatcggtctcaagtca
acccgatgaccagaagactgcaagactggctgaattcggagagatgatacagagccgcttgcacttctctgtcgttgacc
cgctcattgacccgttgaccaaattccatctcttctgctggcgagggaagattacgtggcgccacccacgccaagtgaggag
tgaggtcgcggctgaccaacgagcaagaggtgtttgactttctgcacaattgggcaagtcacacagagccaggtgcaaca
tcagtatggtcaacgggcttttcgcgcacatgggaatgaaggaacagttgaaggcgattatggaagcttcgggtggtggtg
ggggccacggggctggtttgacccaatctggtggcagcaaggtcaacgacagttgttcttgagattctgagtagtcaatacc
gtagaccgcactttcaactgatttctcggtggaaagggttgactaccatgcaattaatcttgcggggtcatttgcgtgacc
ctcgggaggtggtcgagaaattgactggcatagttgacaggttggatgtttgaagagaagtgaagtcacatttgggaatt
ttaactttaaggggtggttaacaattgagcggcattgtcaacgggttggatgctgggaaagtgaatacaacacttggga
gttctgacattgaaggcaagacgtggaattttgatggtgttgaggatatttgatgtgaggttctgatgaattaaagcagg
ggttgatcatttgccagtggaattatgttgggtgaagagagaagggggagaataaacagtggttagagagctatgctgg

Secuencia de aminoácidos predicha de la isoforma nº1 de XylT:

malvnsrgsrvrriakptfvfllinvvcllyffrqnpnpipdaclhgecdkppilvtprwnlkpwpilpsflpwpvsshp
aqqscceayfgnsfnrrtemlkkvegrgwfqclysdtrssvcqgnlrmderirmskggedleevmkreeeeelpkfeeg
sfqiesgygsggeveriatdevldnvvpkgavhvhtmrnlissiqivgphlqcsqwdepvllvtrfeyanlfhtvtdw
ysayassrianlpsrphlifvdghcraeqledmwralfstvrsknfsqpicfrhvvlsplgyetalfkglsefscegap
anrlkvpddqktarlaefgemiraaafdfpvvdpsidpltksilfvrredyvahprhsgrvesrltneqevdfdlhwnash
hrgrcnismvnglfahmgmkeqlkaimeasvvvgahgagtlhlvaarsttvleilssqyrrphqlisrkwgldyhainl
agsfadprevveklgtivdrilgc*

FIG. 3

Alineamiento de secuencia de aminoácidos de 1,2 xiloxiltransferasa de *Lemna minor* y otras plantas superiores

Alineamiento total positivo -- 76,6%
 Identidad total de alineamiento -- 33,7%

Medicago sativa: N° acceso GenBank AY302251.1

Oryza sativa: N° acceso GenBank XM_483221.1

Arabidopsis thaliana: N° acceso GenBank AJ272121.1

Saccharum officinarum: N° acceso GenBank AJ864704.1

Nicotiana tabacum: N° acceso GenBank AJ627182.1

		1		35
Lemna minor XylT	(1)	-----MALVNSRGSRVRR	PTFVFLL	
Medicago sativa XylT	(1)	-----MNNK	YLLF	
Oryza sativa XylT	(1)	MMPVPTYHHHHHNNNSNNH	RRRI	PRVLLAVFAT
Arabidopsis thaliana XylT	(1)	-----MS	RRNP	YLLF
Saccharum officinarum XylT	(1)	---MMAGRAHGHRN	RRRL	PHILLNFAV
Nicotiana tabacum XylT	(1)	-----MNNK	YLLF	
Consenso	(1)		RL RKILKILL	LFAI
		36		70
Lemna minor XylT	(25)	NVYCLLYTERQNFNP	IPDACLHGEC	KPPILVTPR
Medicago sativa XylT	(15)	NSFLCLYTERHH	PSFLNPTPSN	--DIKIFIKD
Oryza sativa XylT	(36)	YAVFAANLRH	QSPHPHPH	PAADPERDAVDAAGG
Arabidopsis thaliana XylT	(19)	NSIFLIYTFVH	SFSFSPEQSQ	---PHIYHVS
Saccharum officinarum XylT	(28)	YVAFATLL	ILQSPQSESHH	QSPDTPRTEARDG
Nicotiana tabacum XylT	(16)	NSIFLYTFESS	SDHFRHKSPQNH	FNTQNHYSLS
Consenso	(36)	NSVSL IYFL HSS		P S
		71		105
Lemna minor XylT	(60)	R-----WNLK	FWILPS	FWVPS
Medicago sativa XylT	(47)	SPFRRPF--IEKHK	CHILKFWILPS	SPWQOPEN
Oryza sativa XylT	(71)	GGGGGAVDRVRVEA	PSSQKFWRLPS	FWTSSEV
Arabidopsis thaliana XylT	(51)	N-----Q	AIQKFWILPS	FWTPFOR
Saccharum officinarum XylT	(63)	VG-----APPYSH	FWRLPS	FWVYSAA
Nicotiana tabacum XylT	(51)	ENHHDNF--HSSVT	QYTFWILPS	FWSQ-NP
Consenso	(71)		S KPWFWILPSYLPWS	SS
		106		140
Lemna minor XylT	(81)	PAQ--GSC	EYFGN	FTTLLKVE
Medicago sativa XylT	(80)	STF-LRSC	EYFGN	FTTLLVDFSR
Oryza sativa XylT	(106)	RPPPKHSC	EYFGN	FTTLLVDFLARG
Arabidopsis thaliana XylT	(74)	NLF-TGSC	EYFGN	FTTLLVDFLARG
Saccharum officinarum XylT	(88)	PP---HTC	EYFGN	FTTLLVDFLARG
Nicotiana tabacum XylT	(83)	NVS-LRSC	EYFGN	FTTLLVDFLARG
Consenso	(106)	P SCEGYFGN	FTTLLVDFLARG	

FIG. 4A

		141		175
Lemna minor XylT	(106)	-----GRGNFOCLMSDTLRSSVCCGGNPMIPERI		
Medicago sativa XylT	(107)	---GRGGGWFRCHSETLRSSICEGGRVPMVVERI		
Oryza sativa XylT	(133)	---GGGGGWFRCHSETLRSSICEGGRVPLIEGLI		
Arabidopsis thaliana XylT	(102)	---GGGGGWFRCHSETLRSSICEGRNIPMVIDRI		
Saccharum officinarum XylT	(112)	---GGGGGWFRCHSETLRSSICEGGRVPMIPALI		
Nicotiana tabacum XylT	(117)	NTVSGGGGWFRCHSETLRSSICEGGRVPMIPALI		
Consenso	(141)	GGGGGWFRCHSETLRSSICEGGRVPMIPALI		
		176		210
Lemna minor XylT	(136)	RMSGGEDLEEVKKEEEELPFFEECSFOIESGY		
Medicago sativa XylT	(139)	KMARGGESLGVIGFEDDEHFLIEDGAEEDGGV		
Oryza sativa XylT	(165)	AMSRGGEPLEQVMGRAPPEELPVEPSALQVEA-A		
Arabidopsis thaliana XylT	(134)	VMSGGGEKLEEVKKEEEELPAPNQGAELVAEEV		
Saccharum officinarum XylT	(144)	AMSRGGEPLEQVMGRAPPEELPVEPSALQVEGPA		
Nicotiana tabacum XylT	(152)	LMSGGGEKLEEVKKEEEELPVEPSALQVEKVTD		
Consenso	(176)	MSRGGGE LEEVMGRAEEEEELPKFE GAFQVEG		
		211		245
Lemna minor XylT	(171)	ESGGE-----VGEKATDEVILNVV		
Medicago sativa XylT	(174)	GFSGG-----KMYDEGTLDRV		
Oryza sativa XylT	(199)	AKETCP-----LVFAGELDAYV		
Arabidopsis thaliana XylT	(169)	SSRLGFKRHRFRGGGEGGSAVSPLVNDENLNEYV		
Saccharum officinarum XylT	(179)	AGKTAP-----LVFAGELDAYV		
Nicotiana tabacum XylT	(187)	KLKIC-----KRLVDEKILNKYL		
Consenso	(211)	A K G		KRLVDDGFLN YV
		246		280
Lemna minor XylT	(191)	PKGAVHVTMRDLIDSIQVGPCHCKSQWIEEPT		
Medicago sativa XylT	(192)	VRGEIMHTMRDLIDSIQVGRKFKDKWIEEPT		
Oryza sativa XylT	(216)	ETGGIGMHMTMRDLIDSGEVVPGGLHCSQWIEEPT		
Arabidopsis thaliana XylT	(204)	QEGGIDHTMRDLIDSIQVGRKFKDKWIEEPT		
Saccharum officinarum XylT	(196)	VRGEIMHTMRDLIDSIQVGRKFKDKWIEEPT		
Nicotiana tabacum XylT	(205)	VRGEIMHTMRDLIDSIQVGRKFKDKWIEEPT		
Consenso	(246)	P GGI RHTMRDLIDSIQVGPGLHCSQWIEEPT		
		281		315
Lemna minor XylT	(226)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Medicago sativa XylT	(227)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Oryza sativa XylT	(251)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Arabidopsis thaliana XylT	(239)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Saccharum officinarum XylT	(231)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Nicotiana tabacum XylT	(240)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
Consenso	(281)	LLVTRFEYANLFHTVTDWYSAYVSSRVTLGLPSRP		
		316		350
Lemna minor XylT	(261)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Medicago sativa XylT	(262)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Oryza sativa XylT	(286)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Arabidopsis thaliana XylT	(274)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Saccharum officinarum XylT	(266)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Nicotiana tabacum XylT	(275)	LVFVDGHCKAQ-LEETWKALFSSVTYAKNFGSPVC		
Consenso	(316)	LVFVDGHCKAQ LEETWKALFSSVTYAKNFGSPVC		

FIG. 4B

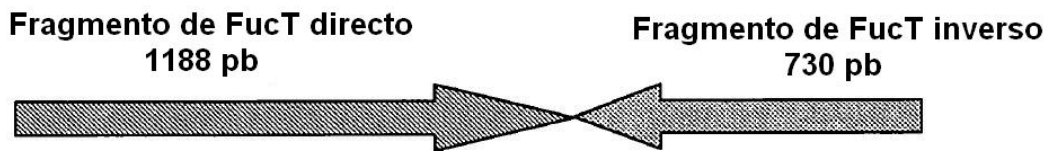
		351		385
Lemna minor XylT	(296)	FRFVLSPLGYETALFKGLSESISCEGASAPANRLKV		
Medicago sativa XylT	(296)	FRHAILSPGLGYETALFKGLTEDIFCDGASAGELNR		
Oryza sativa XylT	(320)	FRHAILSPGLGYETALFKGLSESISCEGASAESLRE		
Arabidopsis thaliana XylT	(308)	FRHAILSPGLGYETALFKGLSGEIDCKGDSAHNLG		
Saccharum officinarum XylT	(300)	FRFVLSPLGYETALFKGLSESISCEGASAQSLRN		
Nicotiana tabacum XylT	(309)	FRHAILSPGLGYETALFKGLTEDIDONGASADILNG		
Consenso	(351)	FRHAILSPGLGYETALFKGLSESISCEGASAQ LWQ		
		386		420
Lemna minor XylT	(331)	NPDDQKTARLSEFGEMIRAAFDFTVDPSIDPLK		
Medicago sativa XylT	(331)	KPDHKTARLSEFGEMIRASFGILINLNHVGPIF		
Oryza sativa XylT	(355)	KPDHKTARLSEFGEMILASFTILRDILSS-ITS		
Arabidopsis thaliana XylT	(343)	NPDDQKTARLSEFGEMIRAAFGILINRHSLEKPL		
Saccharum officinarum XylT	(335)	KPDHKTARLSEFGEMIRASFTILLEDGIVPSKITS		
Nicotiana tabacum XylT	(344)	NPDDQKTARLSEFGEMIRAAFGILINRONIPPTV		
Consenso	(386)	NPDDQKTARLSEFGEMIRAAFGILPVD I KKTS		
		421		455
Lemna minor XylT	(366)	SG-----NLVVRREDYLAHPRHSGKVESRLS		
Medicago sativa XylT	(366)	GG-----HMLFVRREDYLAHPRHSGKVESRLS		
Oryza sativa XylT	(389)	NG-----LNLVVRREDYLAHPRHSGKVESRLS		
Arabidopsis thaliana XylT	(378)	SESSSSASVYNLVRREDYLAHPRHSGKVESRLS		
Saccharum officinarum XylT	(370)	IS-----LFLVVRREDYLAHPRHSGKVESRLS		
Nicotiana tabacum XylT	(379)	G---P-----NLVVRREDYLAHPRHSGKVESRLS		
Consenso	(421)	SG NVLVRREDYLAHPRHSGKVESRLS		
		456		490
Lemna minor XylT	(391)	NEQEVDFLHNWASHHR--G--GKCKINNVNGLFAHM		
Medicago sativa XylT	(394)	NEAEVFESEKSWAAN---YKGTINNVNGLFAHM		
Oryza sativa XylT	(417)	NEKEVYDALEGWAK-----GKCKINNVNGLFAHM		
Arabidopsis thaliana XylT	(413)	NEEEVFDSLHHWATGSTGLTKCGINNVNGLFAHM		
Saccharum officinarum XylT	(398)	NEQEVGAVEKWAK-----GKCKINNVNGLFAHM		
Nicotiana tabacum XylT	(406)	NEEQVFSEKSWALN---HSECKINNVNGLFAHM		
Consenso	(456)	NE EVFDSL WAA G KCKINNVNGLFAHM		
		491		525
Lemna minor XylT	(423)	SMKEQLRAIMASVVIAGHGAGLTHIVSATPTVI		
Medicago sativa XylT	(425)	SMKEQRAIQDASVVIAGHGAGLTHIVSALPKIVI		
Oryza sativa XylT	(447)	NMKEQLRAIQDASVVIAGHGAGLTHIVSATPTDKV		
Arabidopsis thaliana XylT	(448)	SMKEQRAIQDASVVIAGHGAGLTHIVSATNTITI		
Saccharum officinarum XylT	(429)	SMKEQRAIQDASVVIAGHGAGLTHIVSATPTDKV		
Nicotiana tabacum XylT	(437)	SMKEQRAIQDASVVIAGHGAGLTHIVSAEKAIVI		
Consenso	(491)	SMKEQLRAIQDASVVIAGHGAGLTHIVSATP TVI		
		526		560
Lemna minor XylT	(458)	LEIISQYRRPHFALIRWKGLEYHAINLGSYAD		
Medicago sativa XylT	(460)	LEIISQYRRPHFAYIARWKGLEYHAINLDSYAN		
Oryza sativa XylT	(482)	LEIISMYRRPHFALISHWKLEYHAINLPSYAR		
Arabidopsis thaliana XylT	(483)	FEIISVEYRRPHFALIRWKGLEYHAINLANSRAE		
Saccharum officinarum XylT	(464)	LEIISFYRRPHFALISHWKLEYHAINLPSYAS		
Nicotiana tabacum XylT	(472)	LEIISSEYRRPHFALIRWKGLEYHPTYECSYAD		
Consenso	(526)	LEIIS YRRPHFALIRWKGLEYHAINL GSYAD		

FIG. 4C

		561	578
Lemna minor XylT	(493)	PRVVERLEGIDPRIGC-	
Medicago sativa XylT	(495)	ETVINALVSIMKSLRC-	
Oryza sativa XylT	(517)	VTQVNEISNLLGFGC-	
Arabidopsis thaliana XylT	(518)	PTAVESLEIMKSLGC-	
Saccharum officinarum XylT	(499)	IPSTSELSNLLKSLGC-	
Nicotiana tabacum XylT	(507)	PPVTDRLSILPILGC-	
Consenso	(561)	P DVIEKLS ILKSLGC	

FIG. 4D

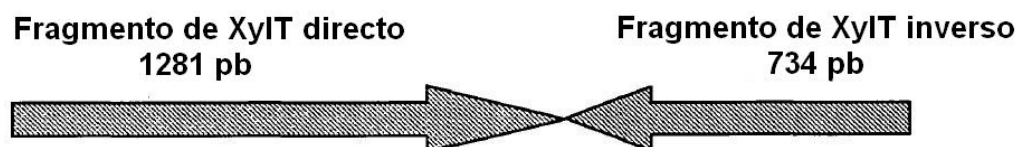
Diseño de ARNi de FucT (bloqueo gen individual):



- La casete de expresión de ARNi de fucosiltransferasa (FucT) se diseñó para expresar un ARN de horquilla con 730 pb de secuencia de ARN de cadena doble y una secuencia espaciadora de 458 pb. La casete de expresión de ARNi comprende un fragmento de FucT directo (nt 255-985 de la SEC ID N°: 1), la secuencia espaciadora (nt 986-1444 de la SEC ID N°: 1) y el fragmento de FucT inverso (versión antisentido de los nt 255-985 de la SEC ID N°: 1).

FIG. 5

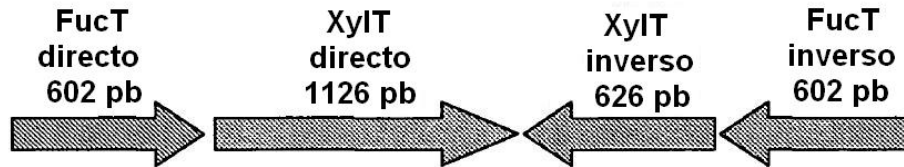
Diseño de ARNi de XylIT (bloqueo de gen individual):



- La casete de expresión de ARNi de xilosiltransferasa (XylIT) se diseñó para expresar un ARN de horquilla con 734 pb de secuencia de ARN de cadena doble y una secuencia espaciadora de 547 pb. La casete de expresión de ARNi comprende el fragmento de XylIT directo (nt 318-1052 de la SEC ID N°: 4), la secuencia espaciadora (nt 1053-1599 de la SEC ID N°: 4) y el fragmento de XylIT invertido (versión antisentido de los nt 318-1052 de la SEC ID N°: 4).

FIG. 6

Diseño de ARNi quimérico de FucT + XylT:



- La casete de expresión de ARNi quimérico de FucT + XylT se diseñó para expresar un ARN de horquilla con 602 pb de secuencia de FucT de cadena doble, 626 pb de secuencia de XylT de cadena doble y una secuencia espaciadora de 500 pb. La casete de expresión de ARNi comprende en la dirección 5' a 3' y ligados operativamente: el fragmento de FucT directo (nt 254-855 de la SEC ID N°: 1), el fragmento de XylT directo (nt 318-943 de la SEC ID N°: 4), una secuencia espaciadora (nt 944-1443 de la SEC ID N°: 4), el fragmento de XylT inverso (versión antisentido de los nt 318-943 de la SEC ID N°: 4) y el fragmento de FucT inverso (versión antisentido de los nt 254-855 de la SEC ID N°: 1).

FIG. 7

Casete de expresión de ARNi de FucT

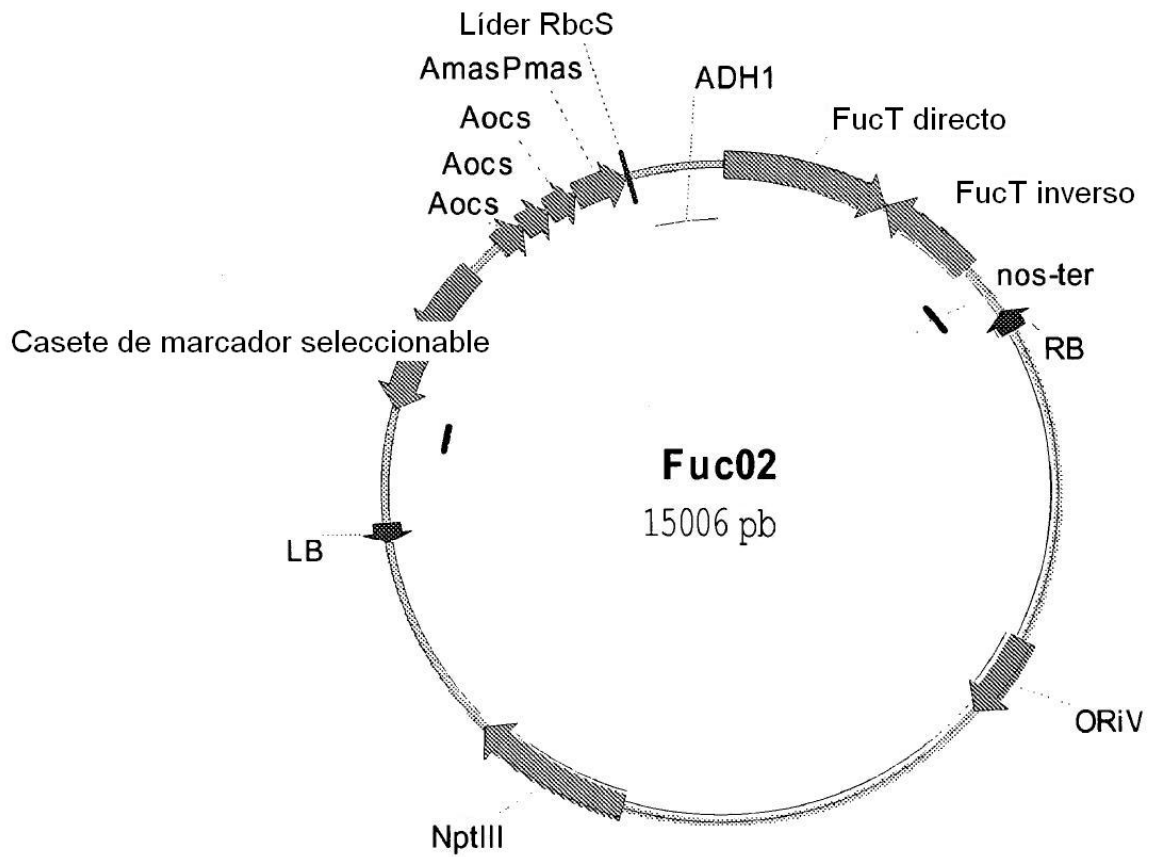


FIG. 8

Casete de expresión de ARNi de XylIT

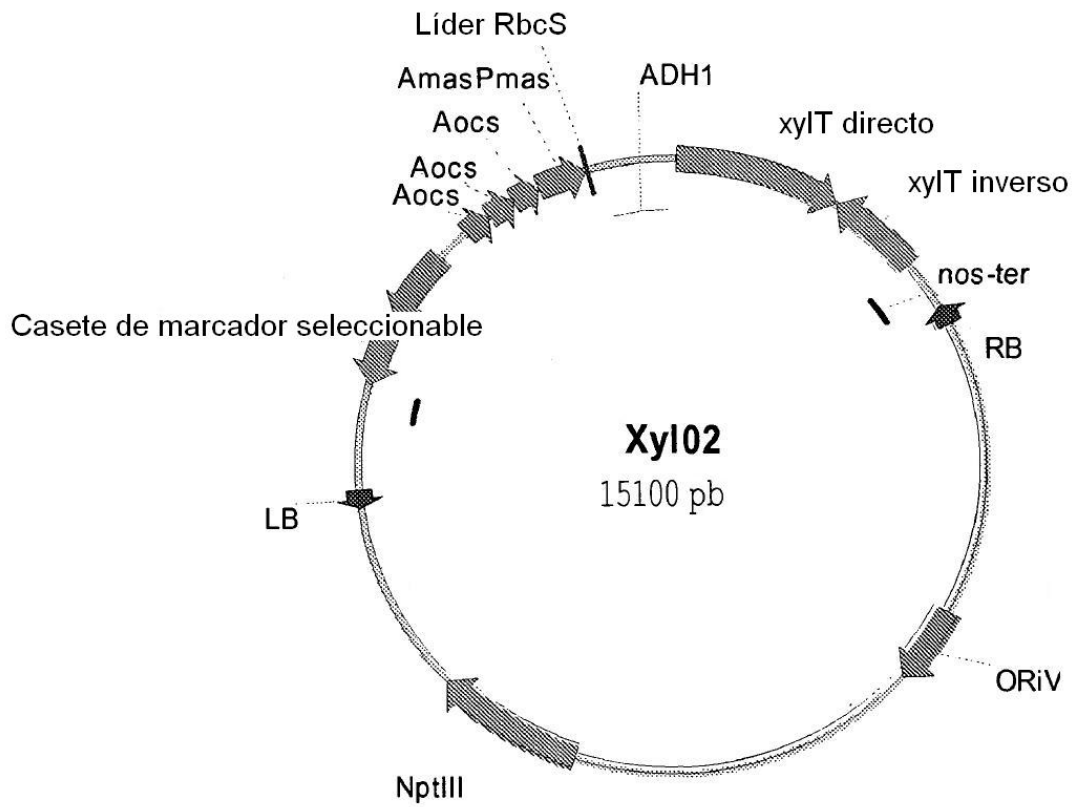


FIG. 9

Casete de expresión de ARNi quimérico de bloqueo doble FucT / XylT

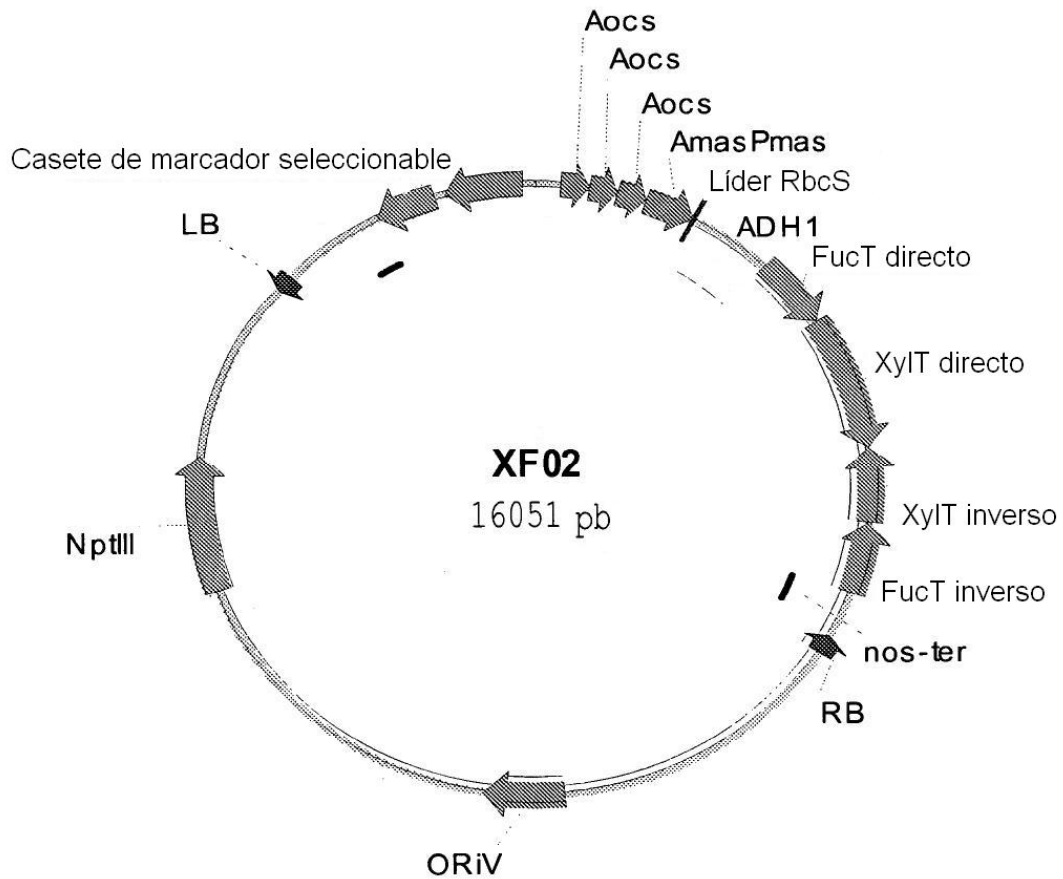


FIG. 10

Casete de expresión de ARNi de gen individual de bloqueo doble de FucT / XylIT

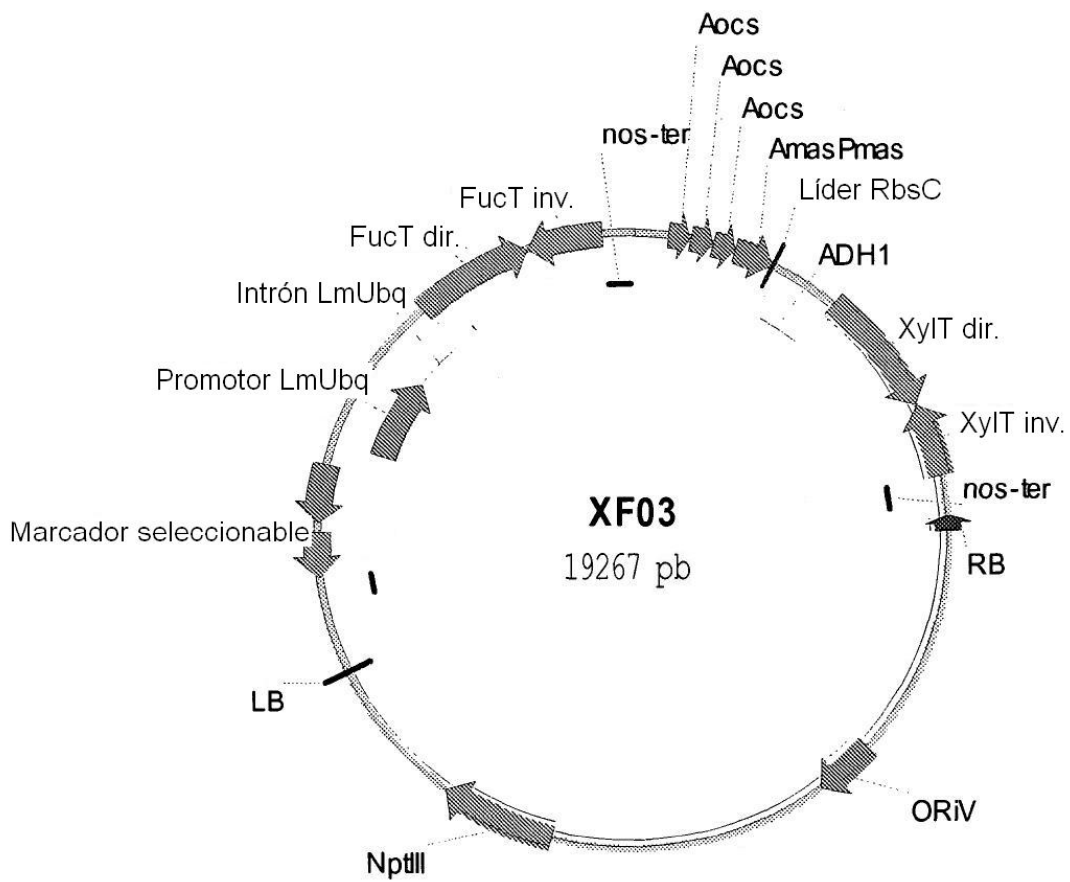


FIG. 11

Co-expresión de IgG1 con bloqueo doble de FucT / XylIT (ARNi Quimérico) N°1

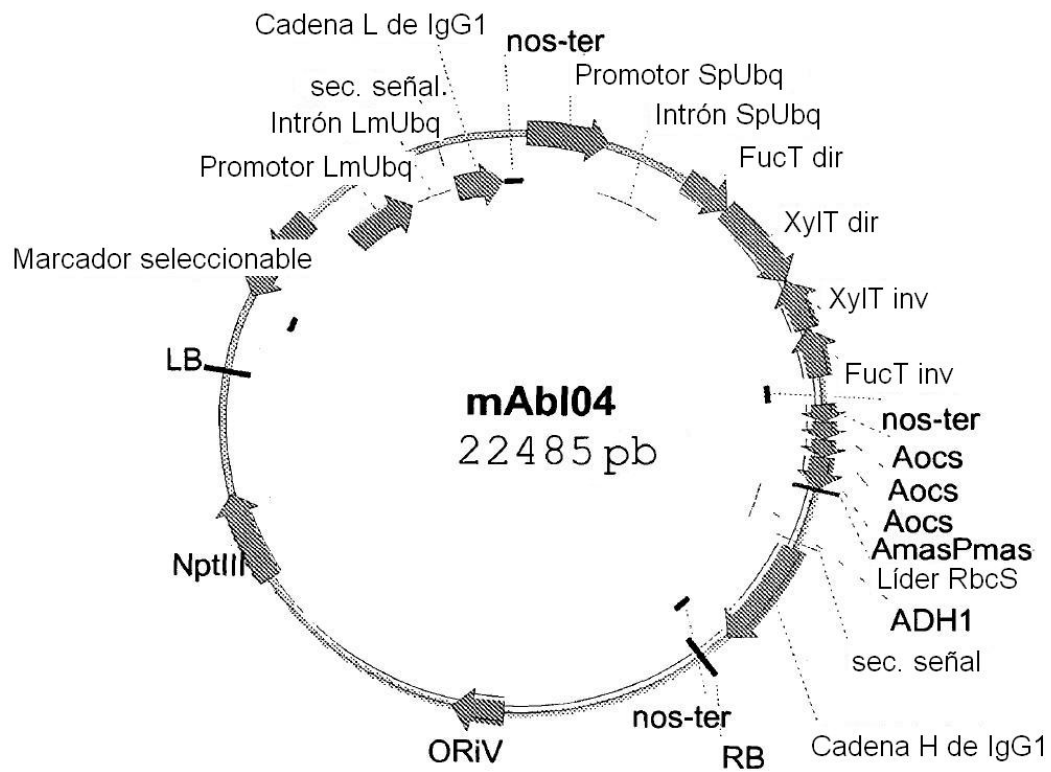


FIG. 12

**Co-expresión de IgG1 con bloqueo doble de
FucT / XylT (ARNi de gen individual)**

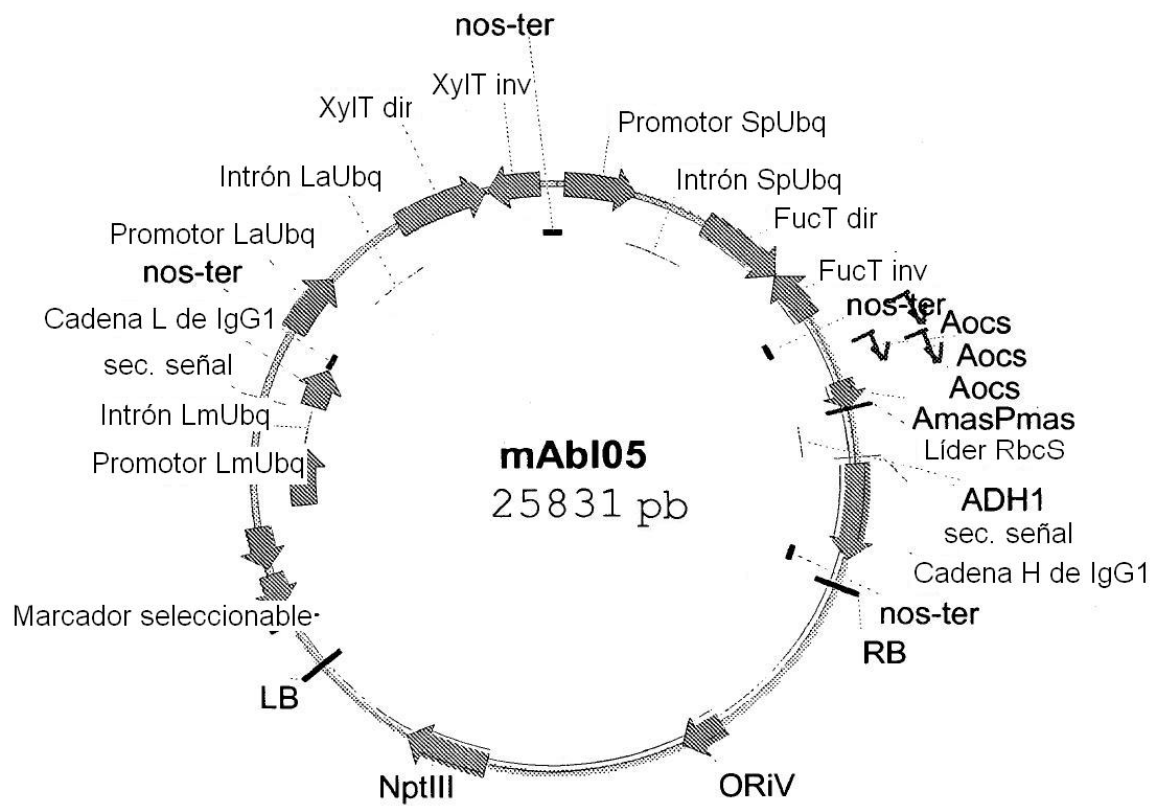


FIG. 13

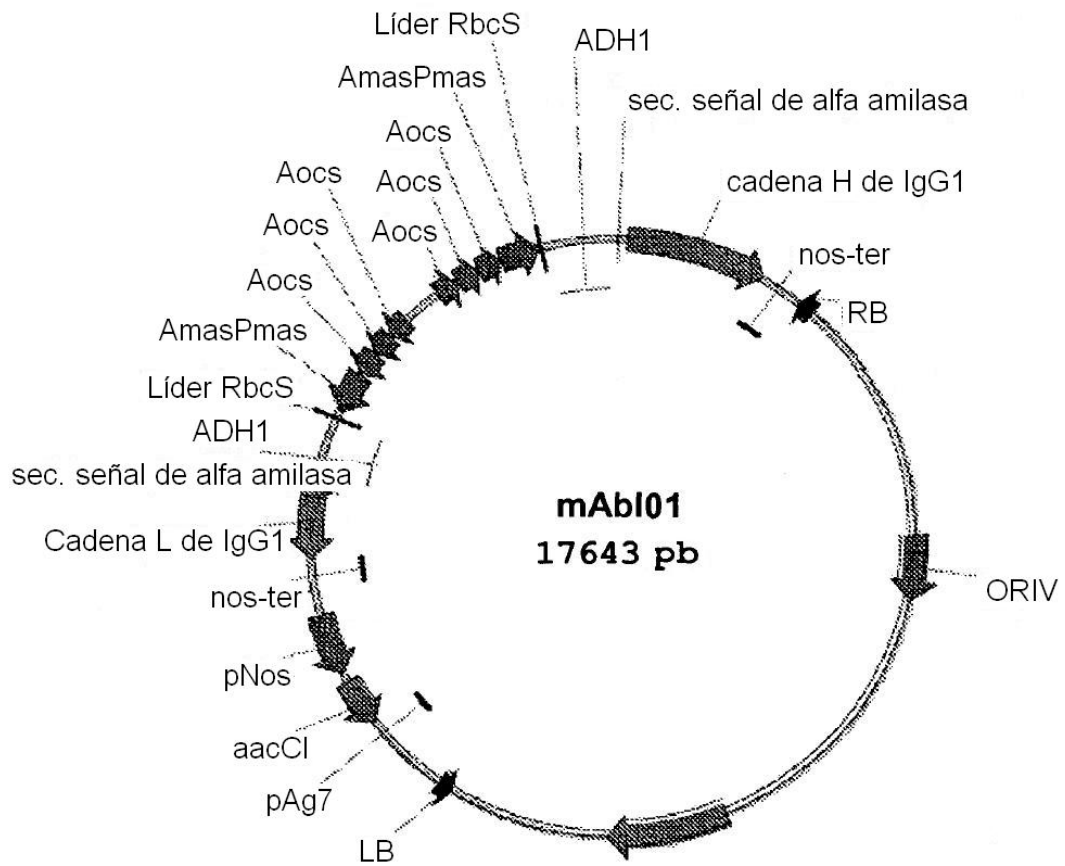
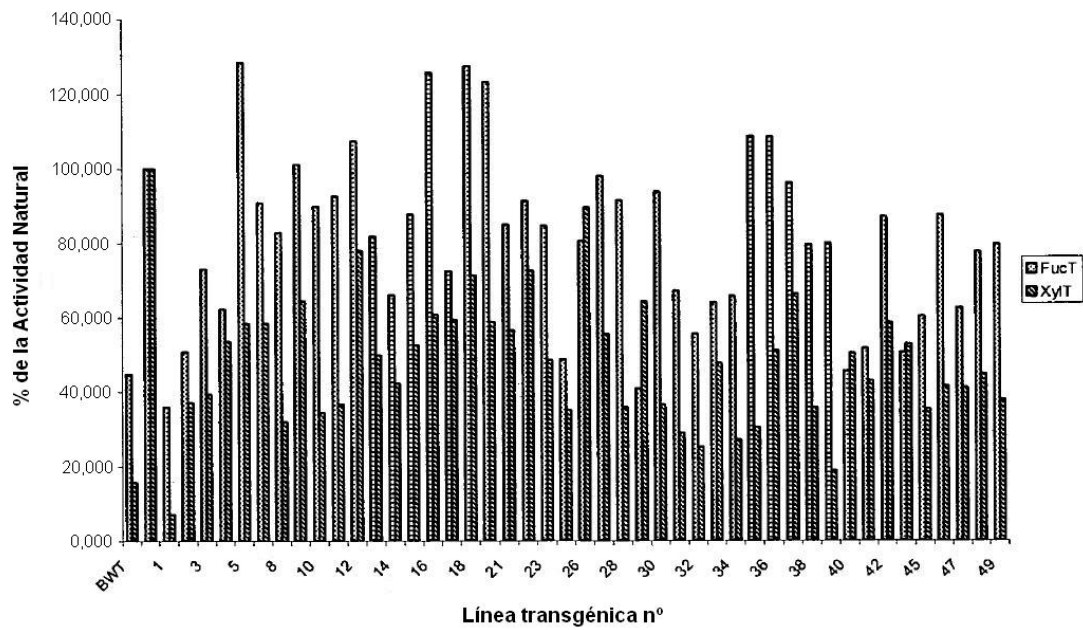


FIG. 14

XF02 Conjunto 1 - escrutinio primario



XF02 Conjunto 2 - escrutinio primario

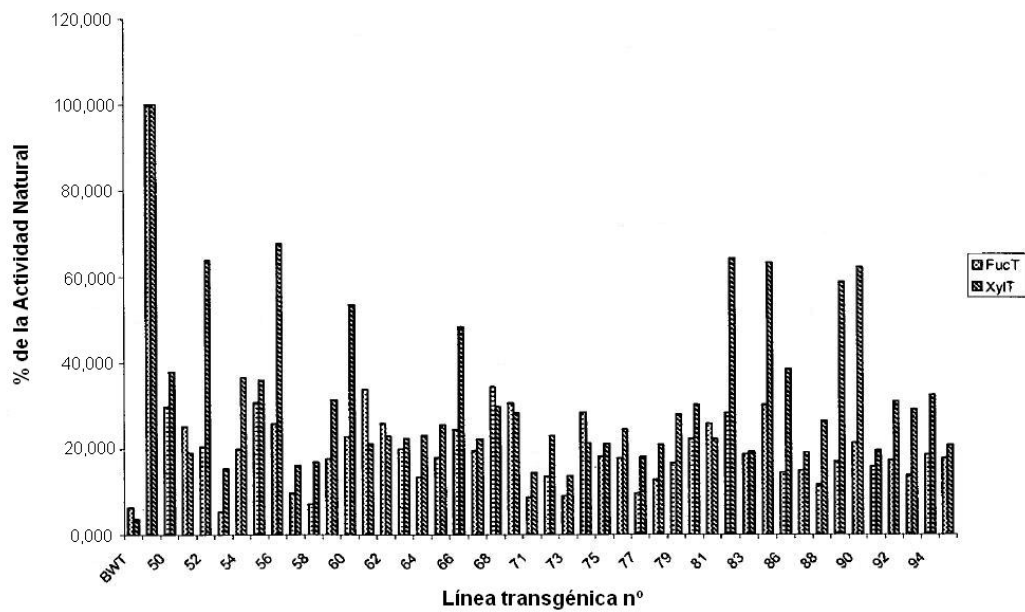


FIG. 15

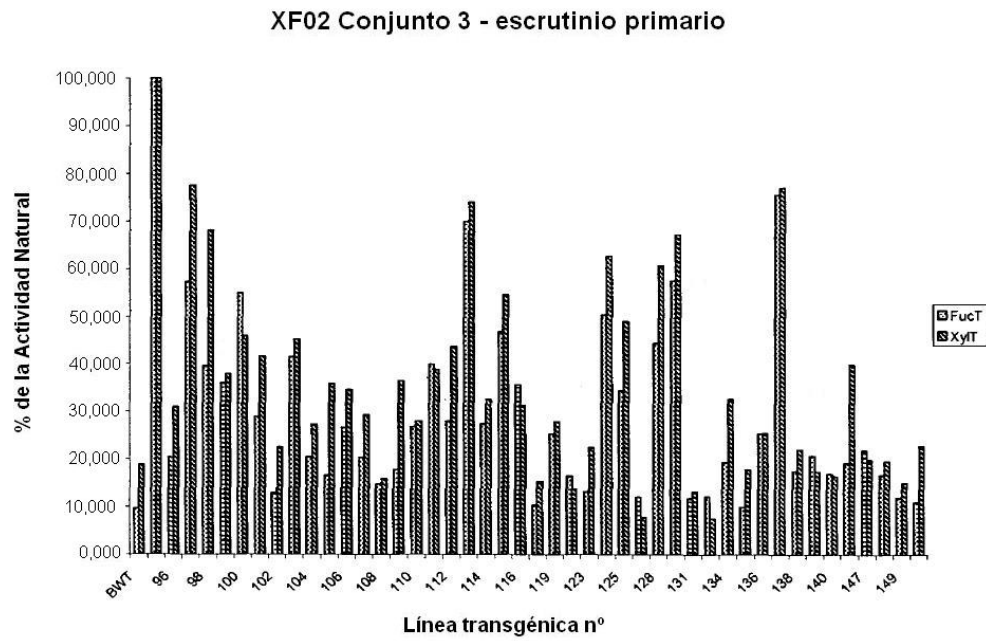


FIG. 16

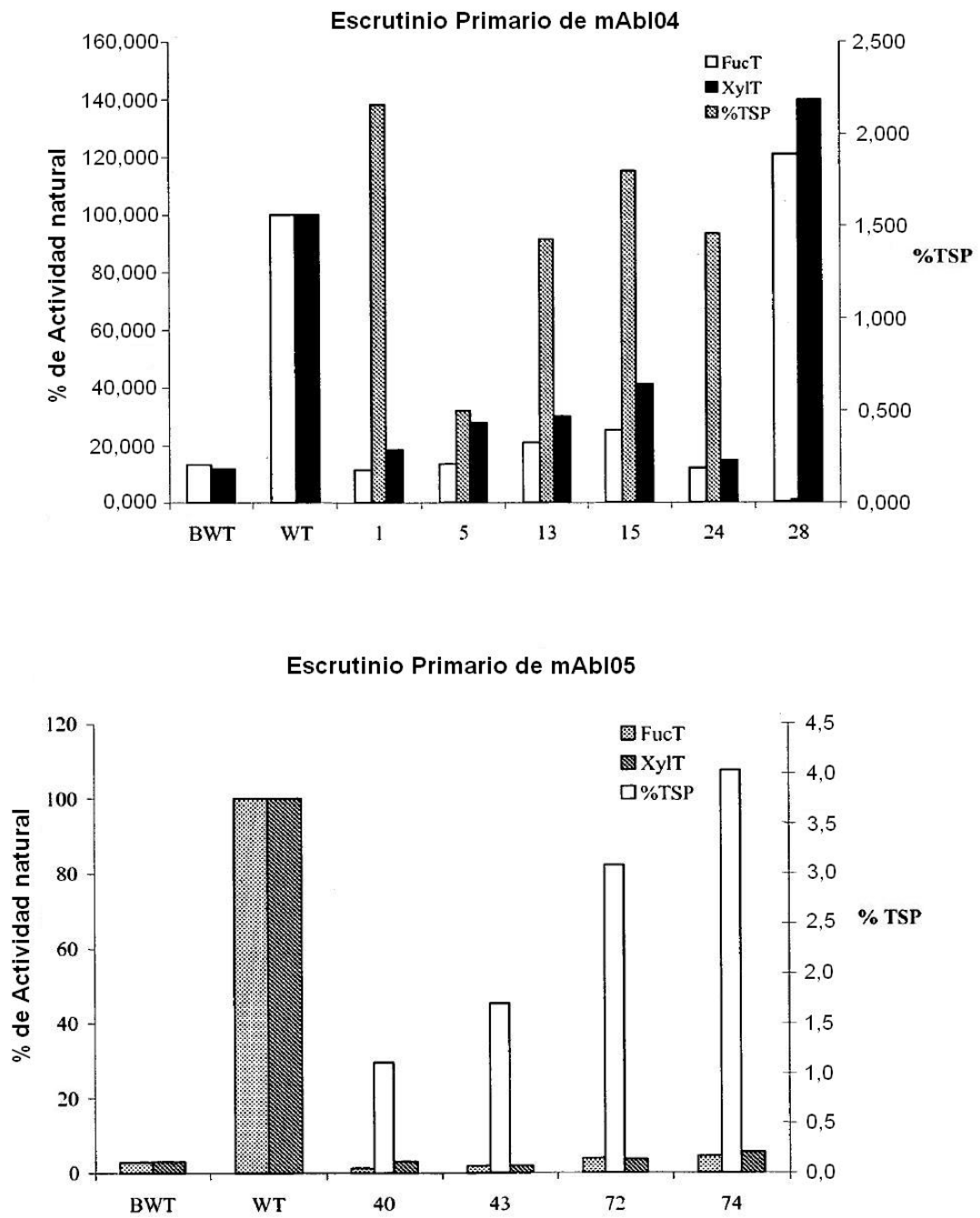


FIG. 17

**Estructura y peso molecular de N-glicanos de mAb de *L. minor*
natural derivatizados**

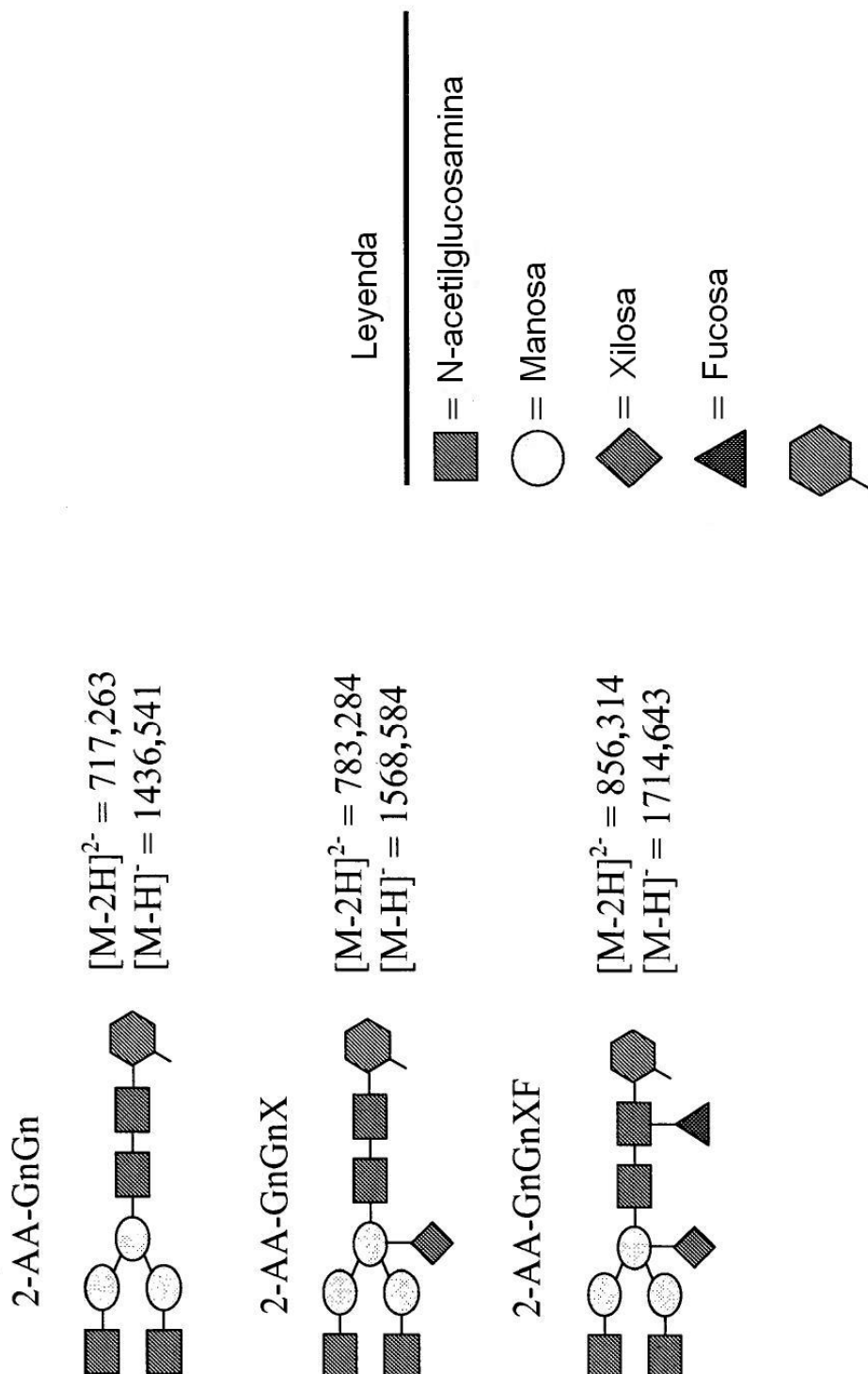


FIG. 18

El mAb101 tiene tres N-glicanos principales

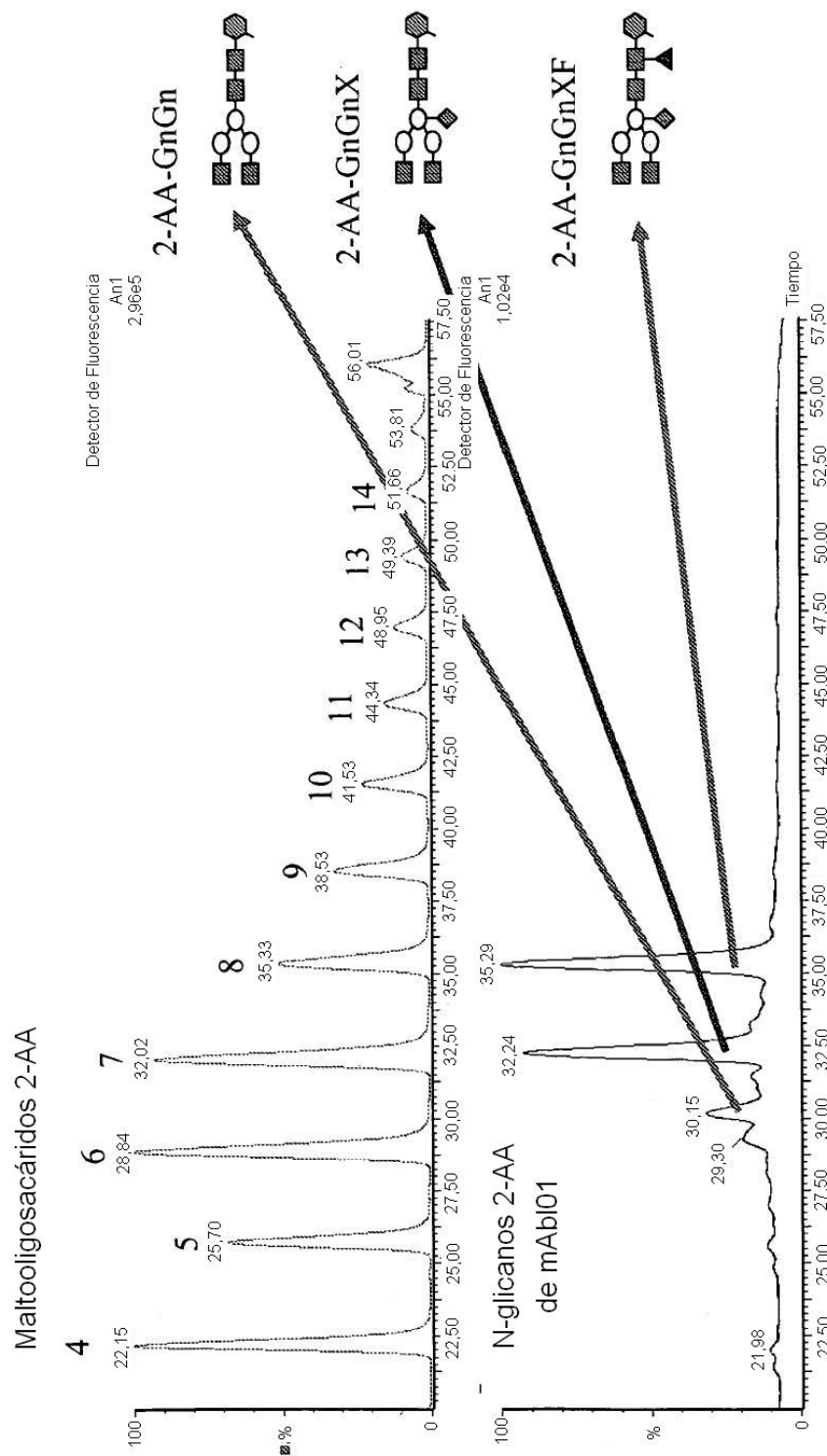


FIG. 19

2-AA-GnGnX

TOF MS ES- 37

$[M-H]^- = 1436,378$

1436,378 37
1437,368 32
1437,410 23
1365,380 11
1227,597 3
1361,781:3
717,694 9
784,197 8
919,281 4
1597,399 8
1821,496 3
2078,734 3
2253,988 3
2394,507 2

2-AA-GnGnX

TOF MS ES- 112

$[M-H]^- = 1568,426$

1568,426 112
1569,388 79
1570,422 39
1742,460 18
1566,402:5
1591,422:5
1840,286:4

2-AA-GnGnXF

TOF MS ES- 118

$[M-2H]^{2-} = 856,721$

856,721 118
857,217 104
857,734 52
858,229 23
690,188 7
790,709 5
1117,063 4
1333,980 4
1712,423:5
1714,478 89
1715,438 69
1716,519 33
1902,380 4

2-AA-GnGnXF

TOF MS ES- 118

$[M-H]^- = 1714,478$

1714,478 89
1715,438 69
1716,519 33
1902,380 4

FIG. 20

El análisis MALDI de N-glicanos de mAbI01 (WT)
confirma los resultados de LC-MS

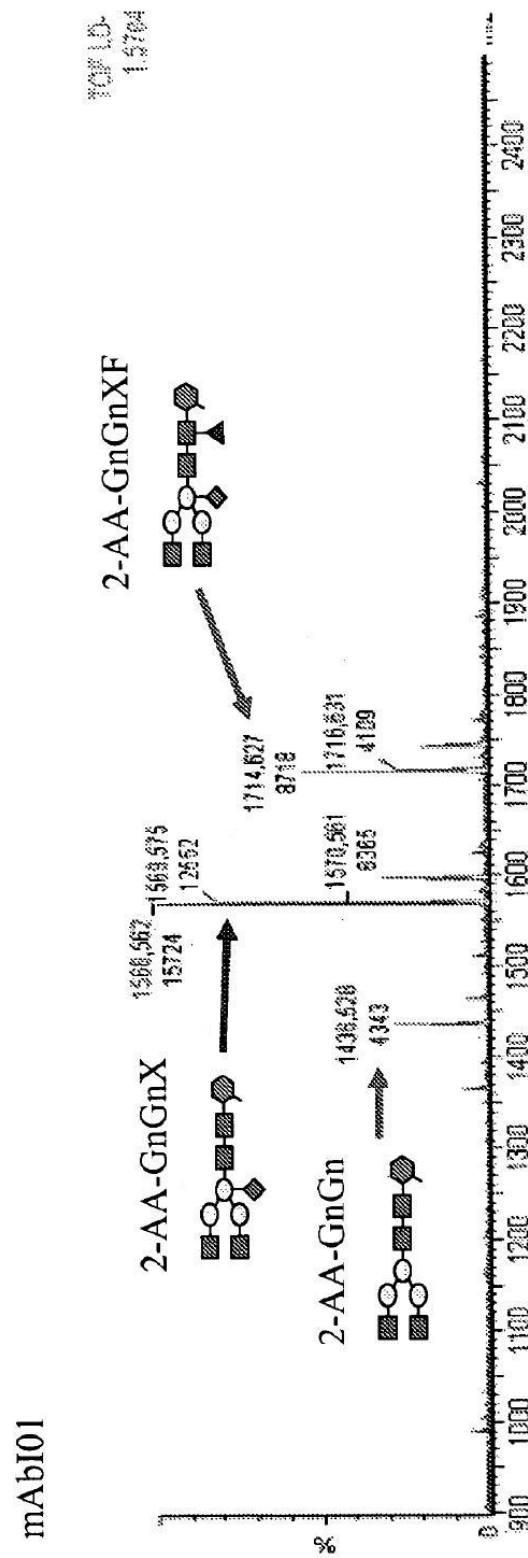


FIG. 21

Superposición de N-glicanos de mAb101 y mAb104

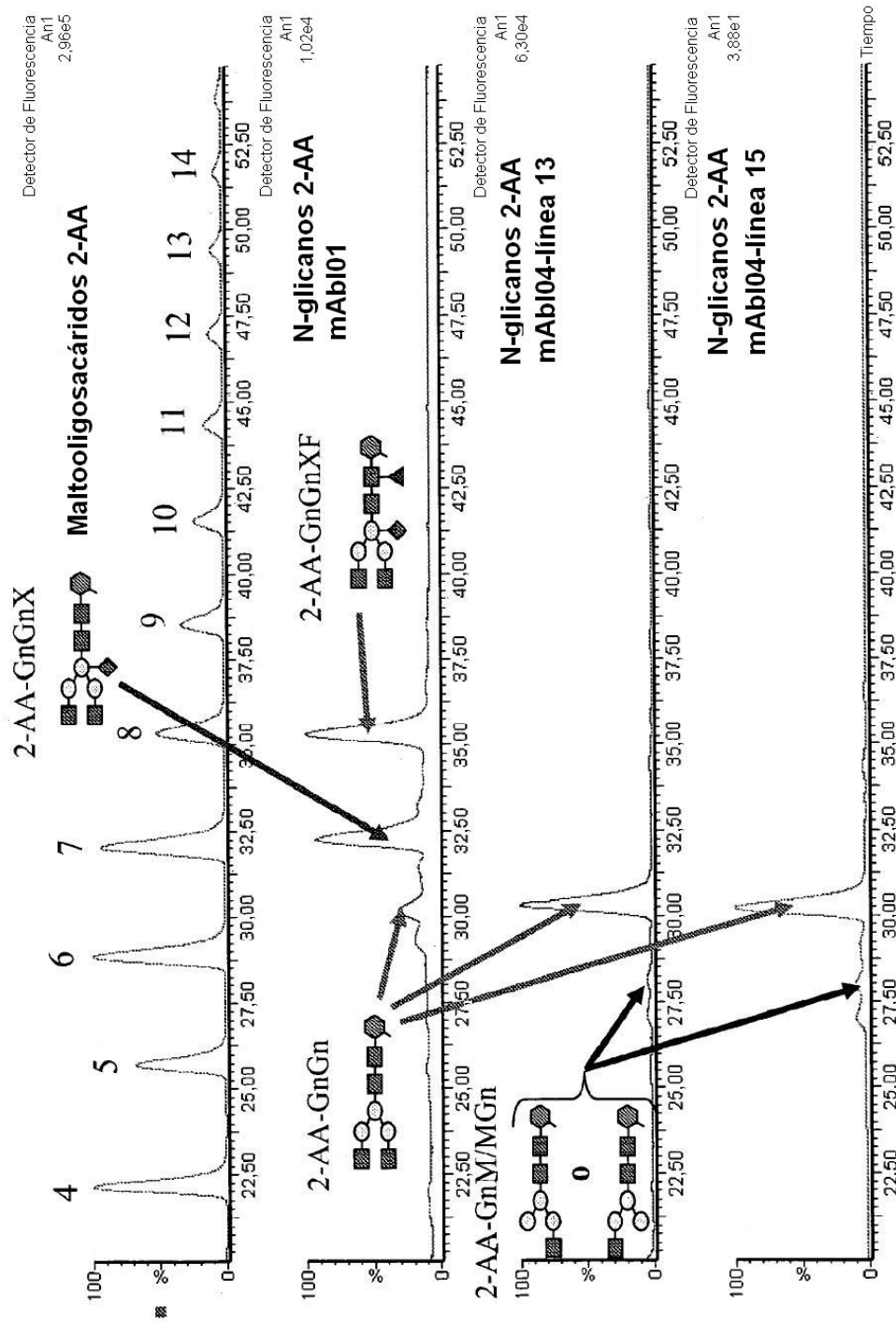


FIG. 22

LC-MS de N-glicanos 2-AA de líneas mAbI04

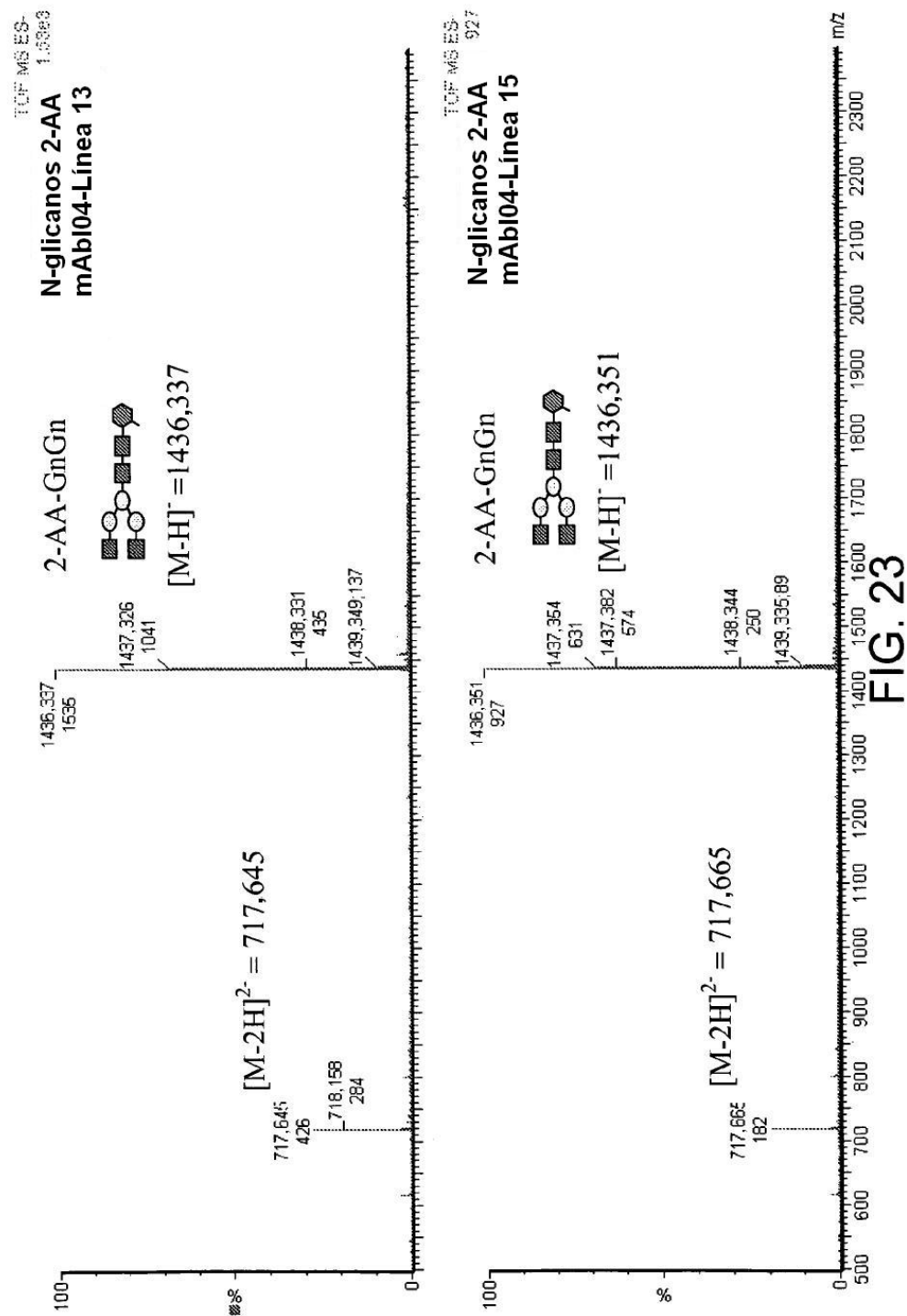


FIG. 23

Análisis MALDI de N-glicanos de mAbI01 y mAbI04

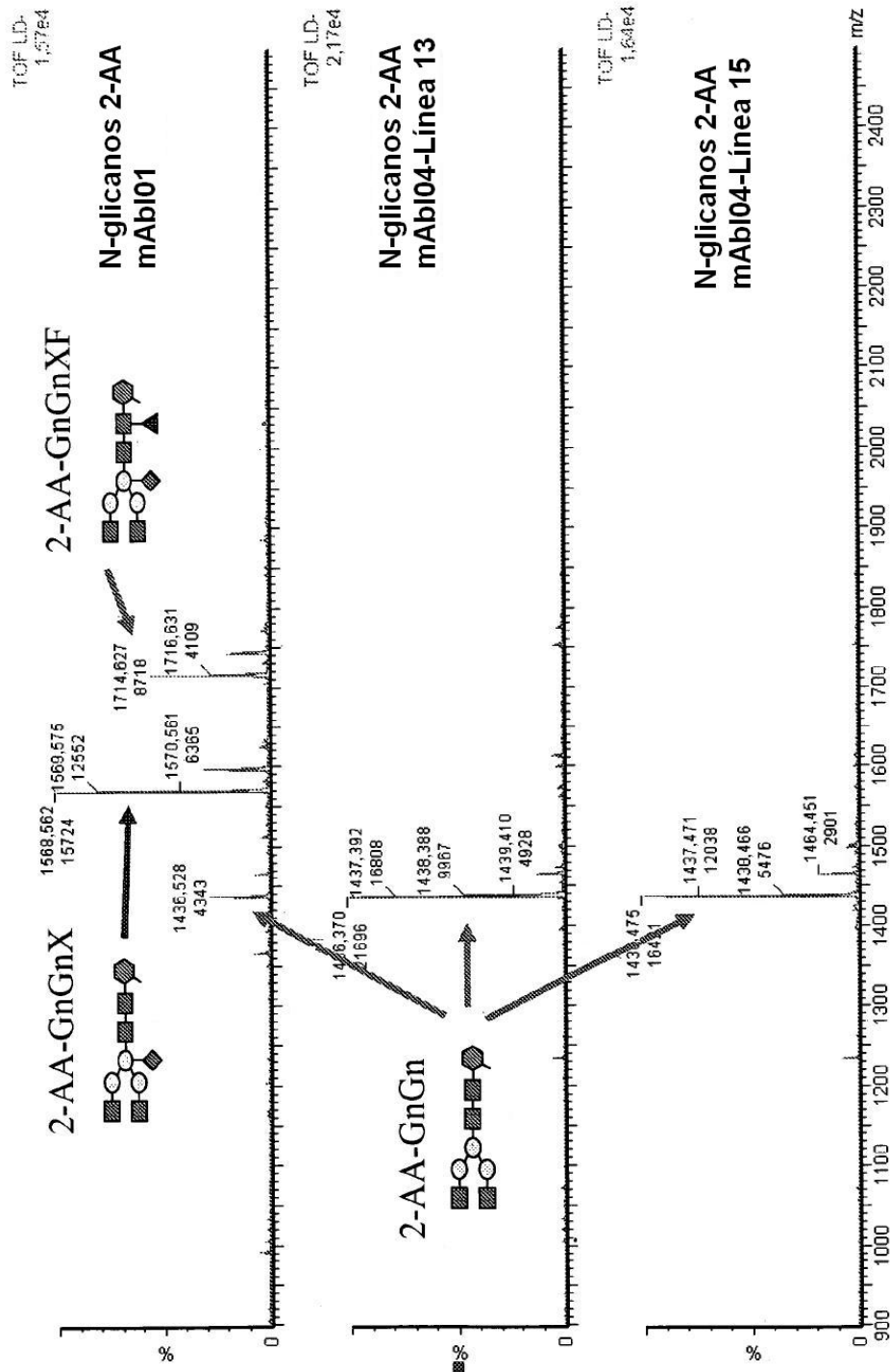


FIG. 24

Análisis de masas intactas de mAb101-20

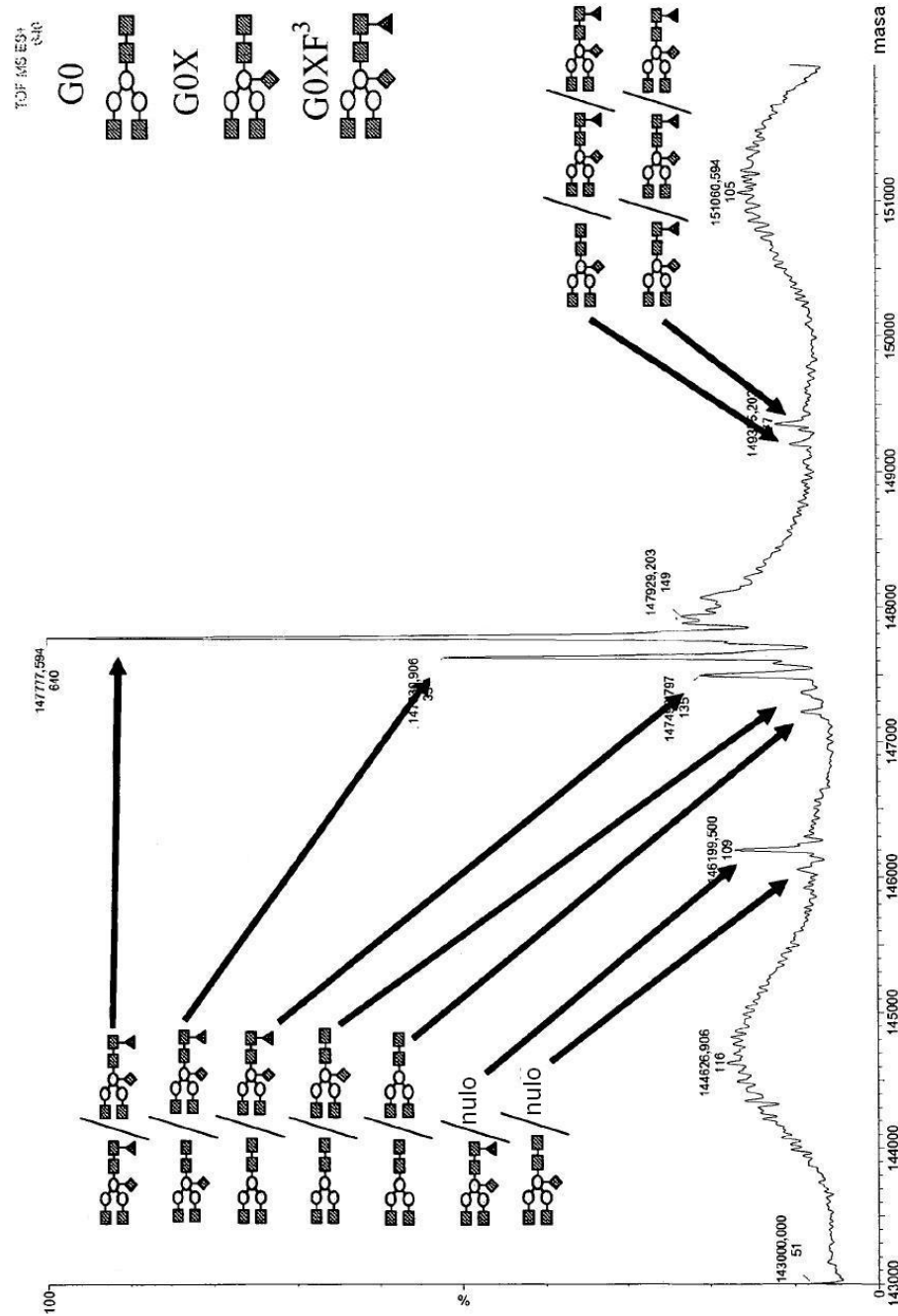


FIG. 25

Análisis de masas intactas de mAbI04-15

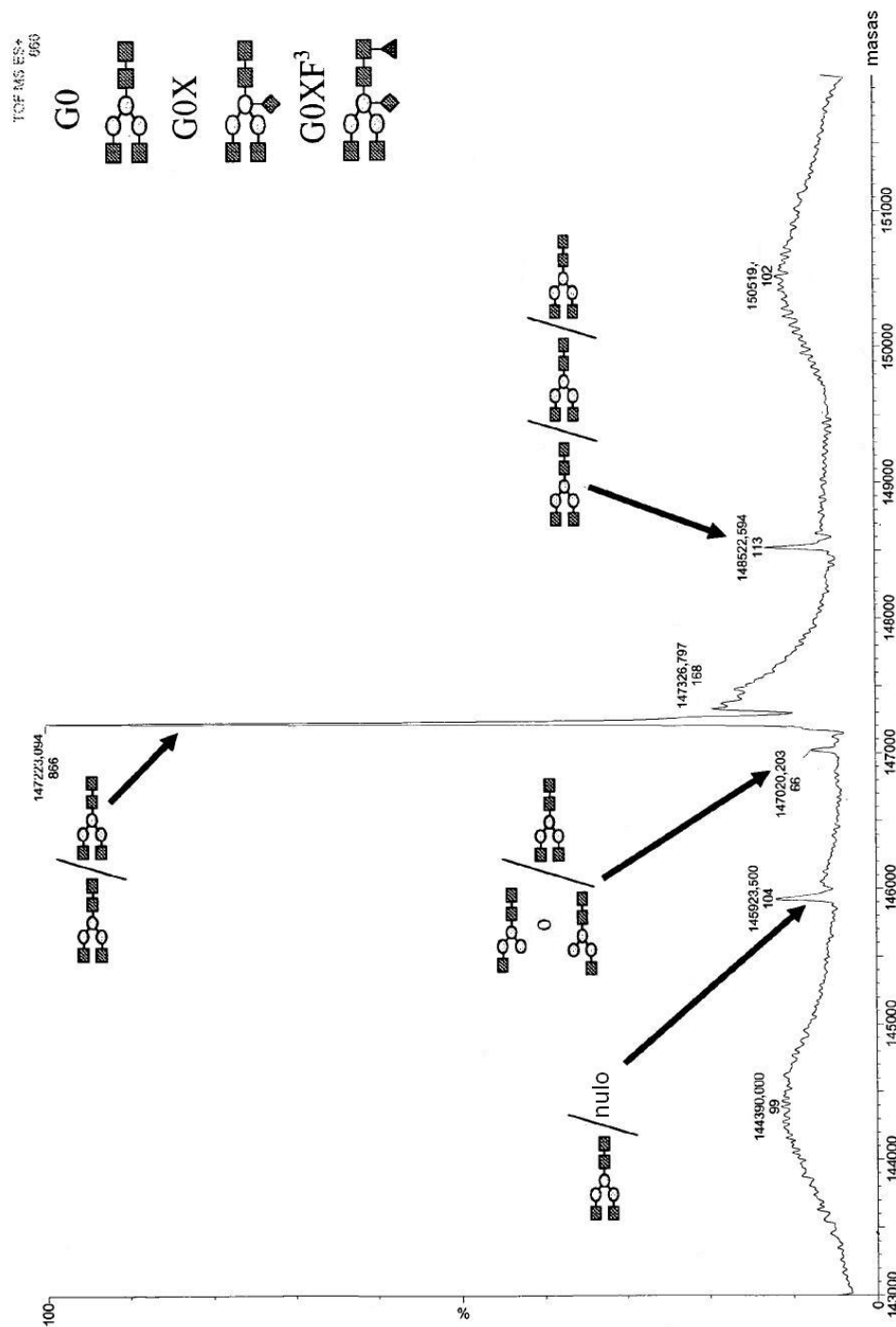


FIG. 26

Análisis de masas intactas de mAbI05-72

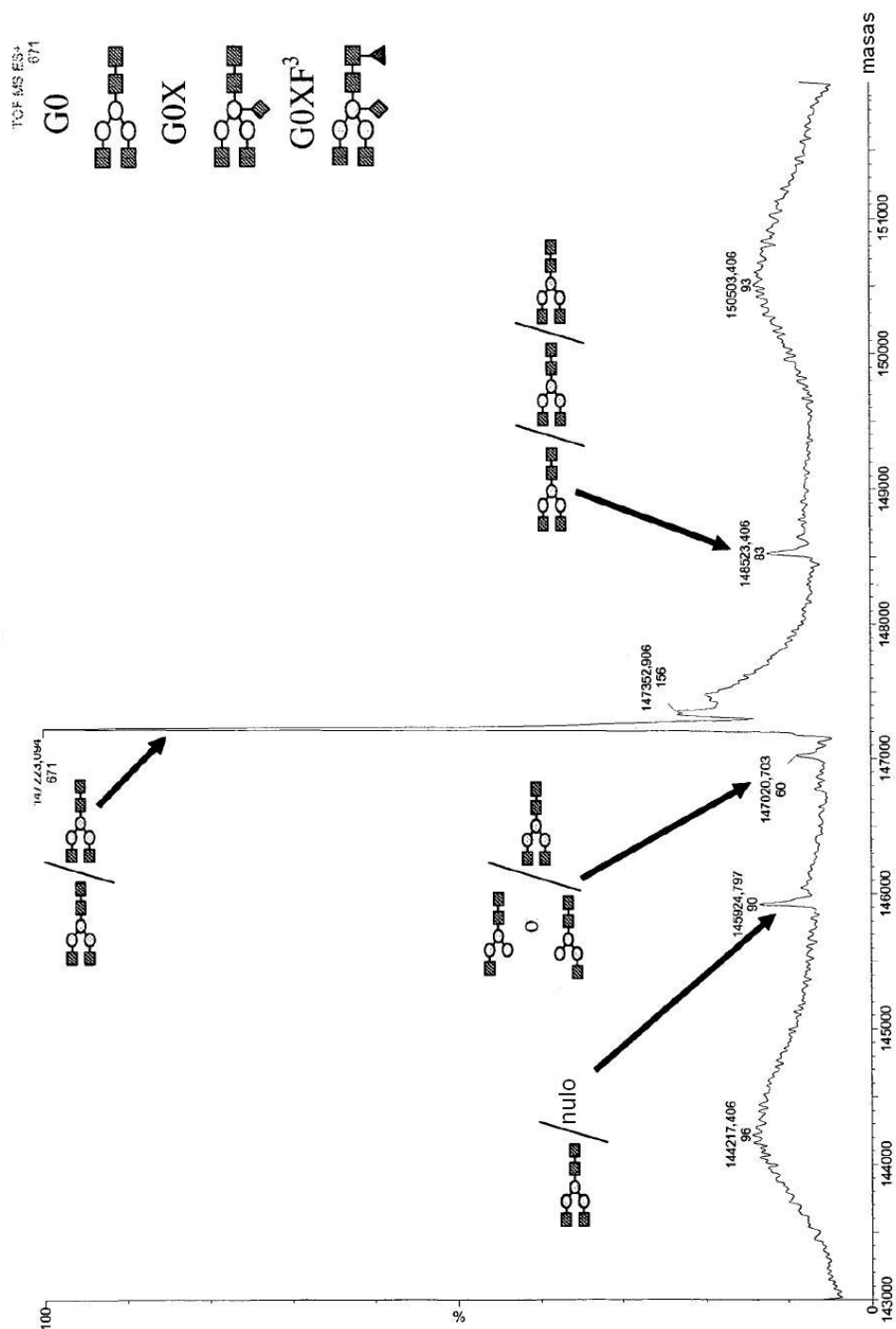
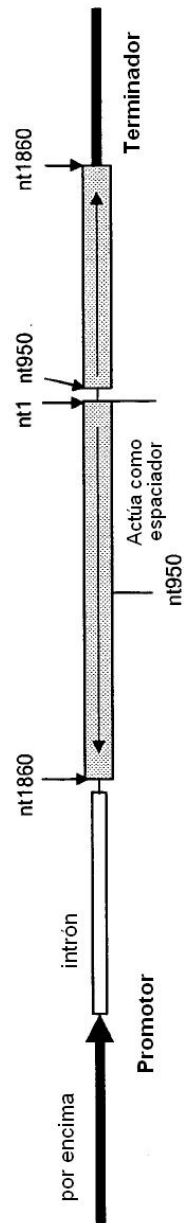


FIG. 27

Possible diseño 1 (para atacar genes individuales; aplicable a ambas secuencias):

Primero clonar la secuencia de longitud completa en la orientación antisentido, a continuación la mitad 3' de la secuencia por debajo de la secuencia de longitud completa pero en sentido orientación. De este modo, la mitad 3' de la secuencia (por ejemplo, los nt 9501860) formarán el tallo de ARNds del ARNhp, y la mitad 5' de la secuencia actuará como espaciador. La región 3' se elige ya que dicha región está relativamente conservada entre diferentes especies de plantas (especialmente para FucT) en comparación con la región 5'.



Possible diseño 2 (para atacar ambos genes):

Fusionar un secuencia de aprox. 500 pb de FucT con una secuencia de aprox. 500 pb de XylT, y entonces usar la secuencia de fusión para preparar las construcciones de ARNhp mostradas a continuación. La secuencia entre nt 700 y nt 1400 representa una región (especialmente para FucT) que está relativamente conservada entre diferentes especies de plantas y por tanto es representativa de una diana potencialmente buena. Ambas versiones mostradas a continuación se pueden preparar para usarse juntas para maximizar potencialmente el silenciamiento ya que los resultados preliminares muestran que la secuencia adyacente al espaciador parece proporcionar un mejor silenciamiento que la secuencia distal de la construcción de ARNhp.

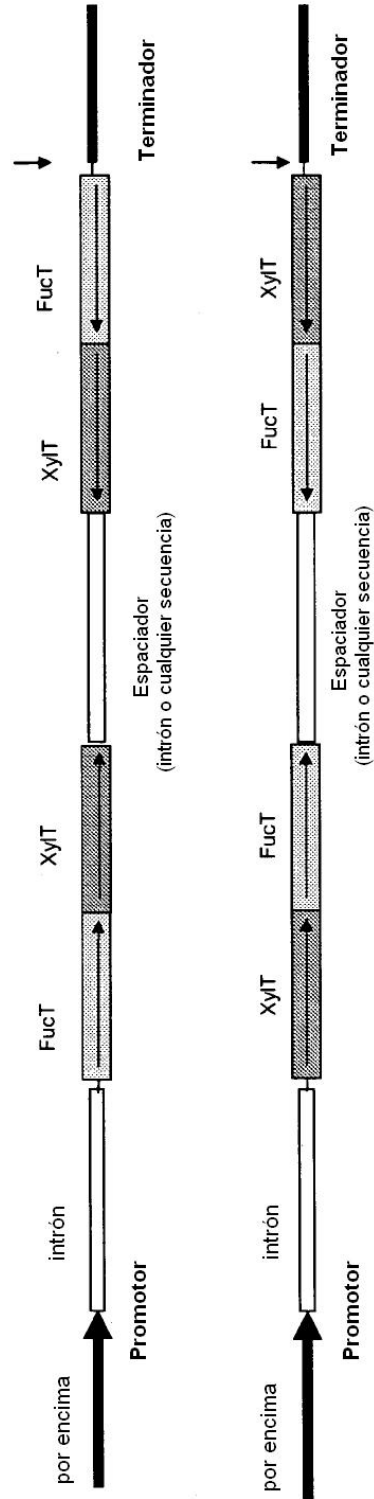


FIG. 28

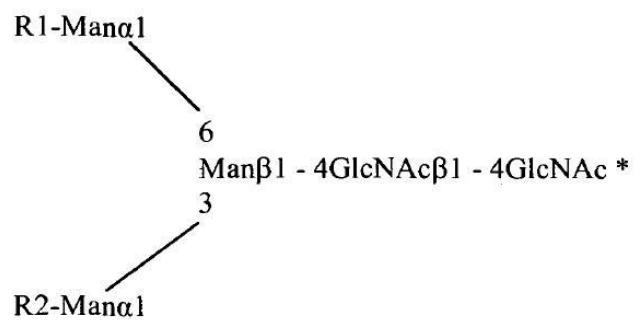


FIG. 29A

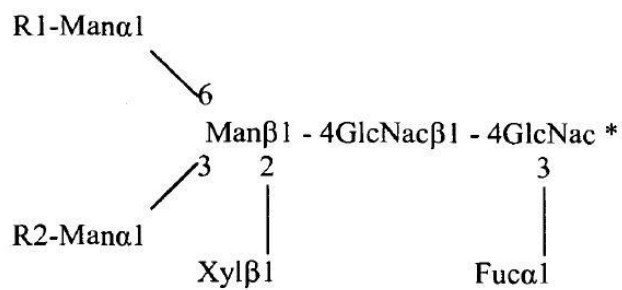


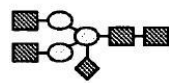
FIG. 29B

ESTRUCTURA DE N-GLICANO

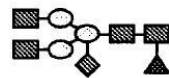
LEYENDA



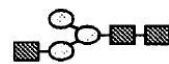
G0 o GnGn o
GlcNAc₂Man₃GlcNAc₂



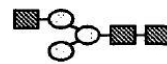
G0X o GnGnX



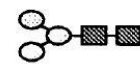
G0XF³ o GnGnXF



MGn o
GlcNAc₁Man₃GlcNAc₂
(unido al brazo de 1,3 Manosa)



GnM o
GlcNAc₁Man₃GlcNAc₂
(unido al brazo de 1,6 Manosa)



Man₃GlcNAc₂

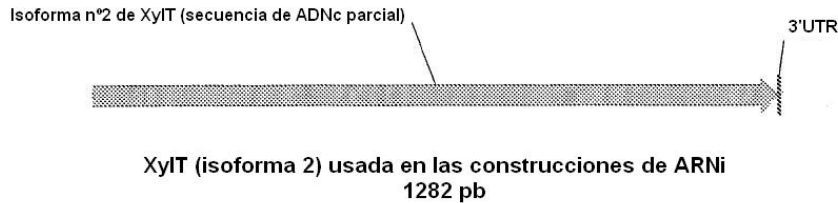
■ N-acetilglucosamina

○ Manosa

◆ Xilosa

▲ Fucosa

FIG. 30

beta 1-2 xilosiltransferasa (isoforma 2) de *Lemna minor*

Secuencia de ADNc parcial de la isoforma n°2 de XylIT (secuencia codificadora en negrita, secuencia 3'UTR en formato estándar):

tgccaagcctacttcggcaacagcttcaaccgcccggacggagatgctgaagaaggtagaggaagaggatggtccagtgccgtgaca
gcgatactcttcgaagttctgtttgccagggagggaatttgcggatggaccggaaaggattaggatgtcgaaaggggggaagatcta
gaggaggtgatgaagagagaggaggaagaagaattgccaaattcgaggaggggtcggtccagattgaatctggttatggaagcggga
ggggaagtggagagagaattgcgactgacgaggtctctcgataatgttgcggaaggcgctgttcattgtacataccatgcgcaatctc
atcagttcgattcagattgttgggtccgggcatcttcaatgctctcagtgagtcgacgaaccgggtctctctgtcaacagcttcgaatacgc
aatctctttcacaccgtcaccgactggtacagccctacgcaagctcgaggattgccaaacttgccctctcgccctcaattgatttctcgat
ggccattgcagggcagaaacagttagaggacagtggtggcgagccctgttctcaaccgtccgatacggcaagaacttctccagccagctctgc
ttccgccacgcgtctctctccctcttggctatgagacagctctcttcaaggcctatcagagagcttcagctgtgagggagtgccggccaa
tcagctcaaagtcaaccctgatgaccagaagactcgagactggctgaattcggagagatgatcagggtgcctttgacttctctgtctt
gacccggcgttgacccgttgaccaaattccatctcttctgcggcggaagattacgtggcgacccacggcacagtgggagagtgga
tcgctgttgaccaatgagcaagaggtgtttgacttctgcacaaatgggcaagtcaacacagaagcaggtgcaacgtcagttgtgtcaac
gggttttcgcgacatgggaatgaaggaaacaggtgaaggcaattatggaagcttcggtggtgtcggggccacggggctggtttgact
catctgttgcgacgaagggtcaacgacagttgttcttgagattctgagcagtcataatcgtagaccgcacttcaactgatttcacggtgaa
aggttggactaccacgaattaatcttccgggtcgtatcgtgactctcgggaggtggtcgagaaattgactggcatagtcgatggctt
ggatttgaagataag

Secuencia de aminoácidos predicha de la isoforma n°2 de XylIT (secuencia parcial):

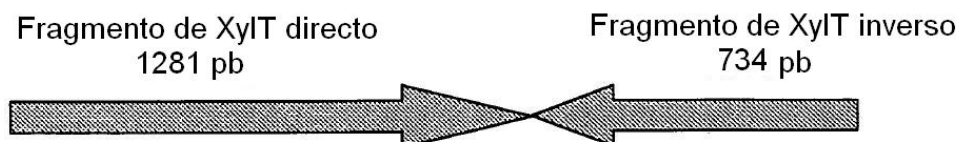
ceayfgnsfnrrtemlkkvegrgwfqclysdtrssvcqgnlrmddperimrskggdleevmkreecelpkfeegsfqiesgygsggev
geriatdevldnvpkgavhvhtmrnlissiqivgphlqcsqwdepvllvtrfeyanlfhtvtdwysayassrianlpsrphlifvdghcrae
qledtwralfstvryaknfsqpvcfhavlsplgyetalfklsesfscgvpnanqlkvnppddqktarlaefgemiraafdfpvpdppvdpplks
ilfvrredyvahprhsgrvesrltneqevdfhlkwasqhrsrcnsvvnglfahmgmkeqvkaimeasvvvgahgagltlhaarsttvl
eilssqyrphfqlisrkwgldyhainlagsyadprevveklgtgvdglgc*

FIG. 31

Isoforma nº1 de LmXylT	(1)	1	40
Isoforma nº2 de LmXylT	(1)	-----	
Isoforma nº1 de LmXylT	(41)	41	80
Isoforma nº2 de LmXylT	(1)	-----	
Isoforma nº1 de LmXylT	(81)	81	120
Isoforma nº2 de LmXylT	(1)	-----	
Isoforma nº1 de LmXylT	(121)	121	160
Isoforma nº2 de LmXylT	(36)	-----	
Isoforma nº1 de LmXylT	(161)	161	200
Isoforma nº2 de LmXylT	(76)	-----	
Isoforma nº1 de LmXylT	(201)	201	240
Isoforma nº2 de LmXylT	(116)	-----	
Isoforma nº1 de LmXylT	(241)	241	280
Isoforma nº2 de LmXylT	(156)	-----	
Isoforma nº1 de LmXylT	(281)	281	320
Isoforma nº2 de LmXylT	(196)	-----	
Isoforma nº1 de LmXylT	(321)	321	360
Isoforma nº2 de LmXylT	(236)	-----	
Isoforma nº1 de LmXylT	(361)	361	400
Isoforma nº2 de LmXylT	(276)	-----	
Isoforma nº1 de LmXylT	(401)	401	440
Isoforma nº2 de LmXylT	(316)	-----	
Isoforma nº1 de LmXylT	(441)	441	480
Isoforma nº2 de LmXylT	(356)	-----	
Isoforma nº1 de LmXylT	(481)	481	510
Isoforma nº2 de LmXylT	(396)	-----	

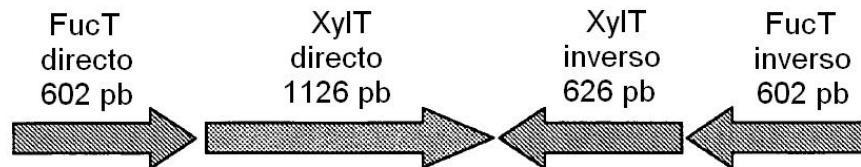
FIG. 32

Diseño de ARNi de XylT (bloqueo de gen individual):



- La casete de expresión de ARNi de xilosiltransferasa (XylT) se diseñó para expresar un ARN de horquilla con 734 pb de secuencia de ARN de cadena doble y una secuencia espaciadora de 547 pb, basada en la secuencia de ADN parcial de la isoforma nº2 de XylT. La casete de expresión de ARNi comprende el fragmento de XylT directo (nt 1-734 de la SEC ID N°: 19), la secuencia espaciadora (nt 735-1282 de la SEC ID N°: 19) y el fragmento de XylT inverso (versión antisentido de los nt 1-734 de la SEC ID N°: 19).

FIG. 33

Diseño de ARNi quimérico de FucT + XylT:

- La casete de expresión de ARNi quimérico de FucT + XylT se diseñó para expresar un ARN de horquilla con 602 pb de secuencia de FucT de cadena doble, 626 pb de secuencia de XylT de cadena doble, y una secuencia espaciadora de 500 pb. La casete de ARNi abarca en la dirección 5' a 3' y ligados operativamente: el fragmento de FucT directo (nt 254-855 de la SEC ID N°: 1), el fragmento de XylT directo de la isoforma n°2 (nt 1-626 de la SEC ID N°: 19), una secuencia espaciadora (nt 627-1126 de la SEC ID N°: 19), el fragmento de XylT inverso (versión antisentido de los nt 1-626 de la SEC ID N°: 19) y el fragmento de FucT inverso (versión antisentido de los nt 254-855 de la SEC ID N°: 1).

FIG. 34

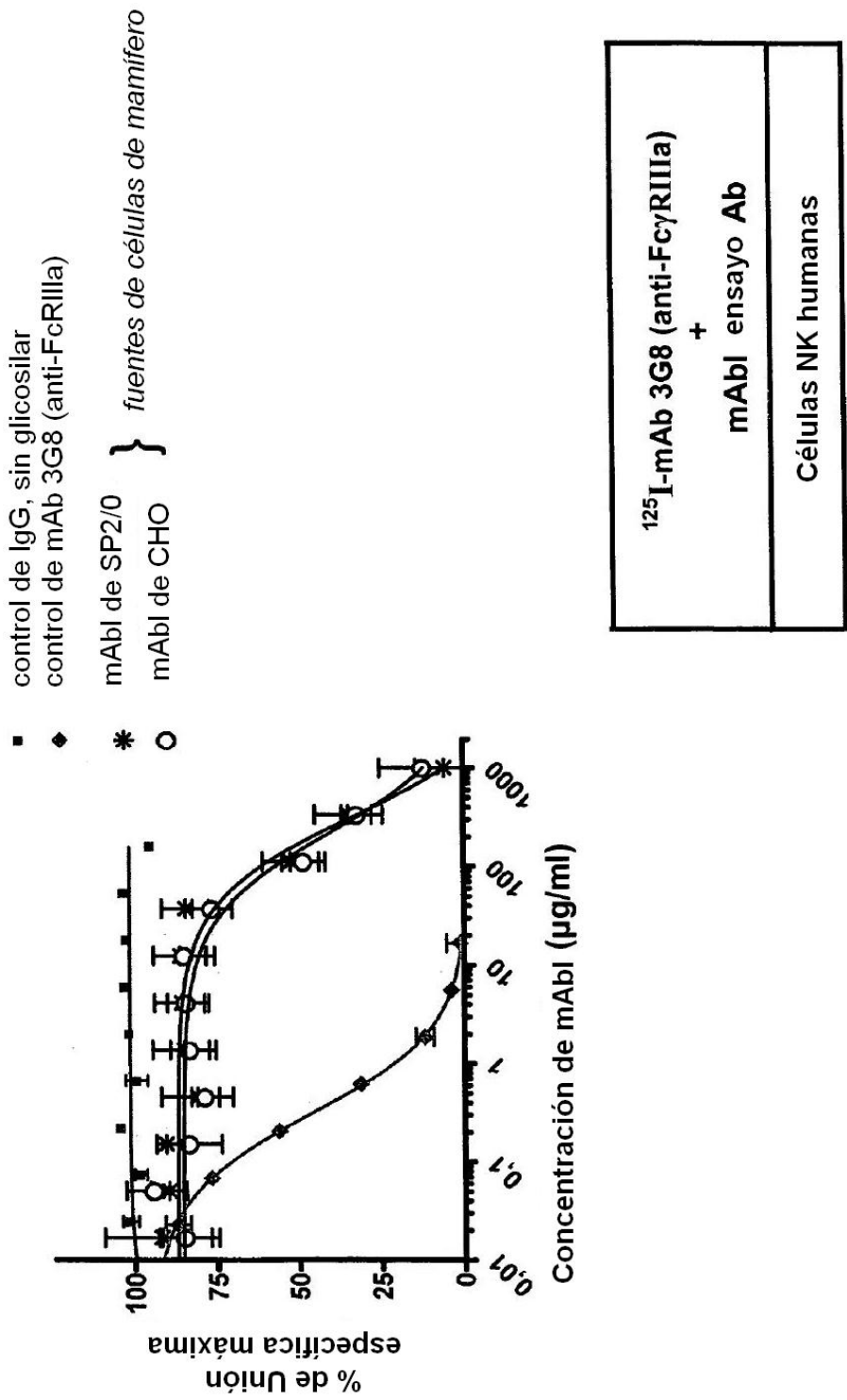


FIG. 35

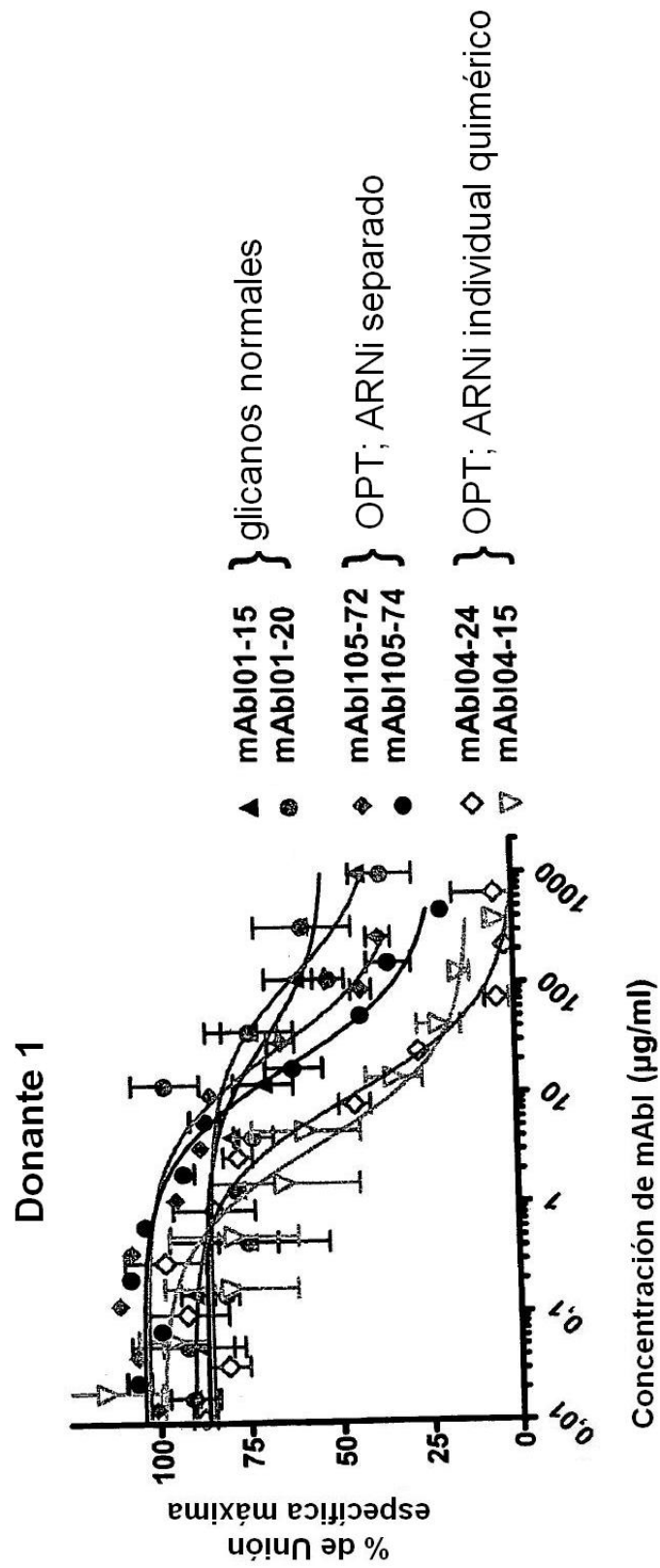
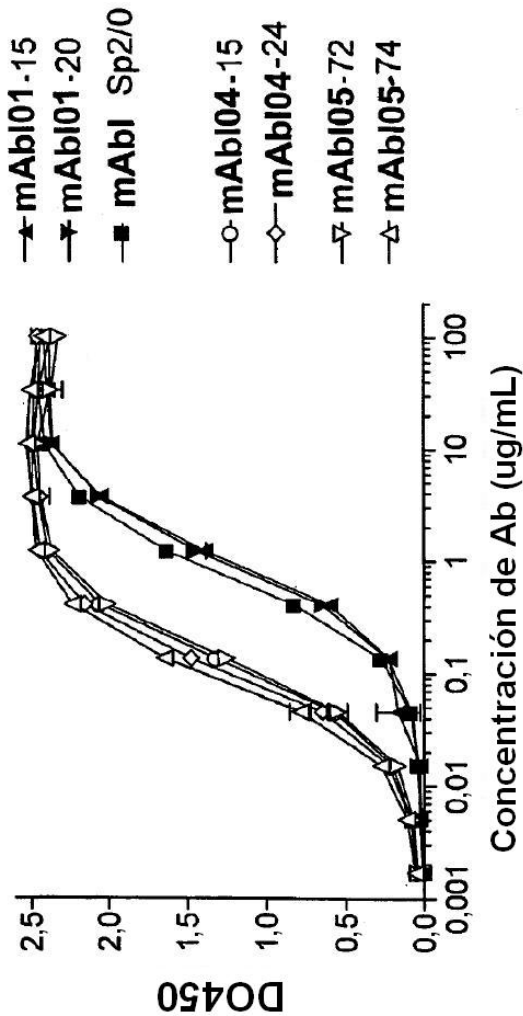


FIG. 36

Unión a FcγRIV de ratón (fucosa-sensible)



Fab anti-humano de cabra F(ab') ₂
muestra mAbI
His-FcγRIV
Cu+

FIG. 38

Vector de Expresión MDXA04

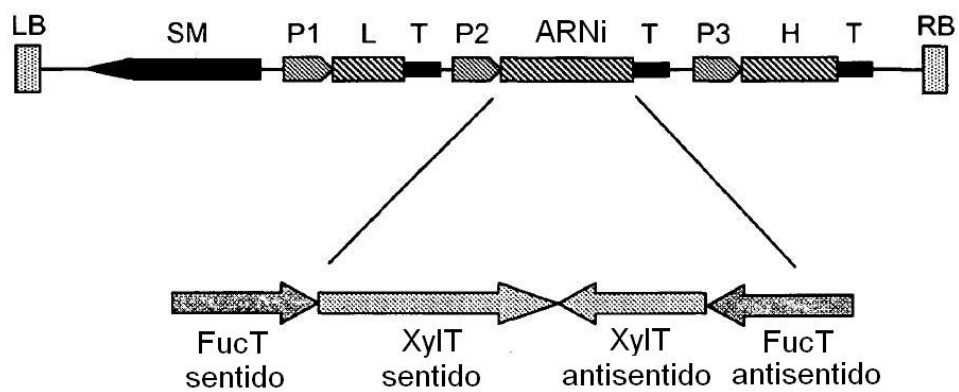


FIG. 39

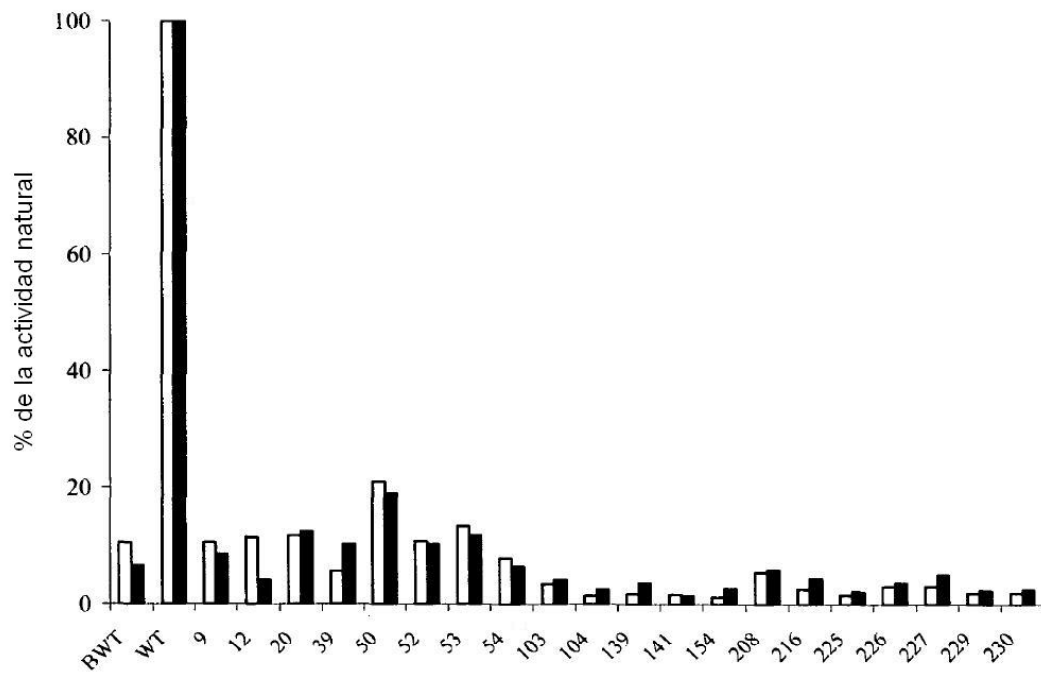


FIG. 40

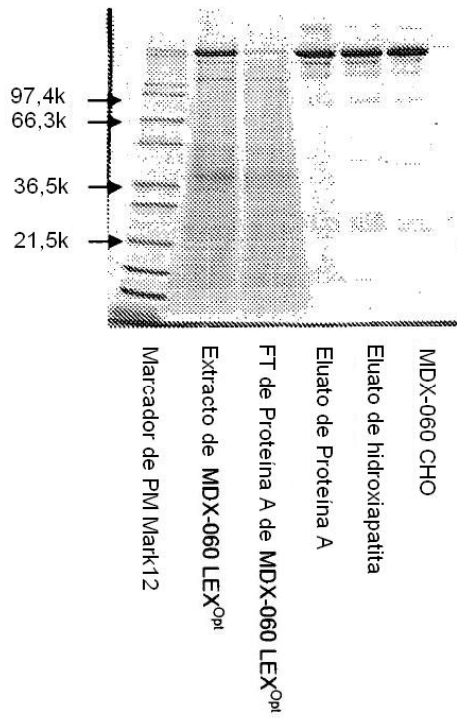


FIG. 41A

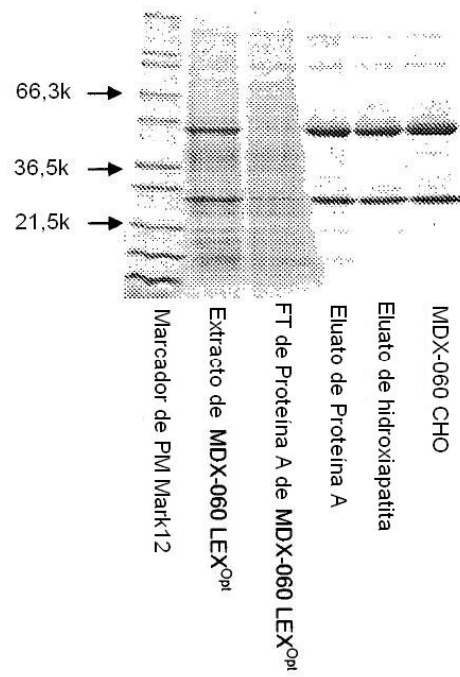


FIG. 41B

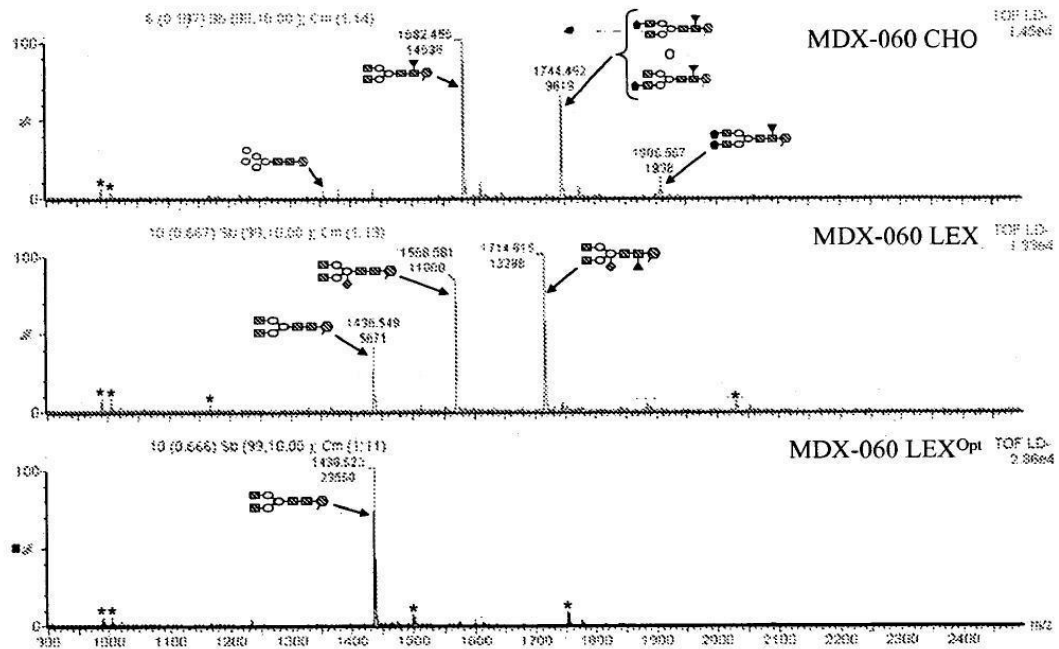


FIG. 42

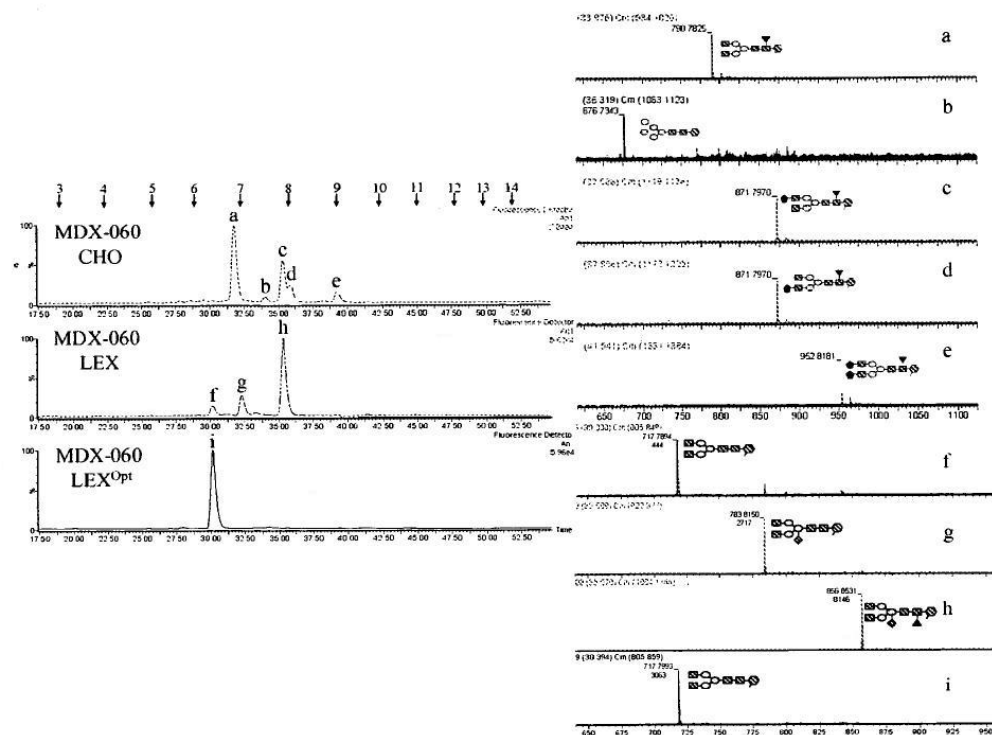


FIG. 43

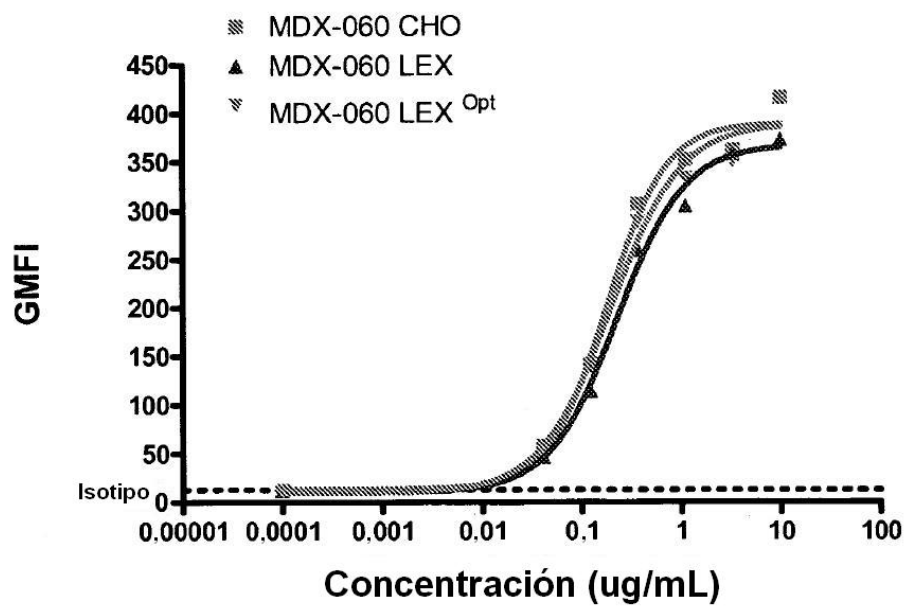
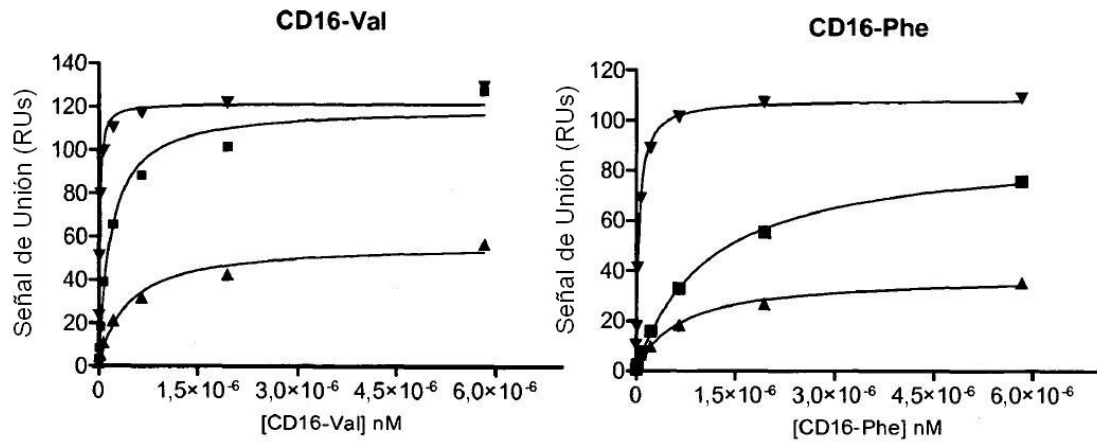


FIG. 44



	Constantes de Unión (nM)		
FcR Humano	MDX-060 CHO	MDX-060 LEX	MDX-060 LEX ^{Opt}
CD16-Val	177	427	12
CD16-Phe	1091	641	41

FIG. 45

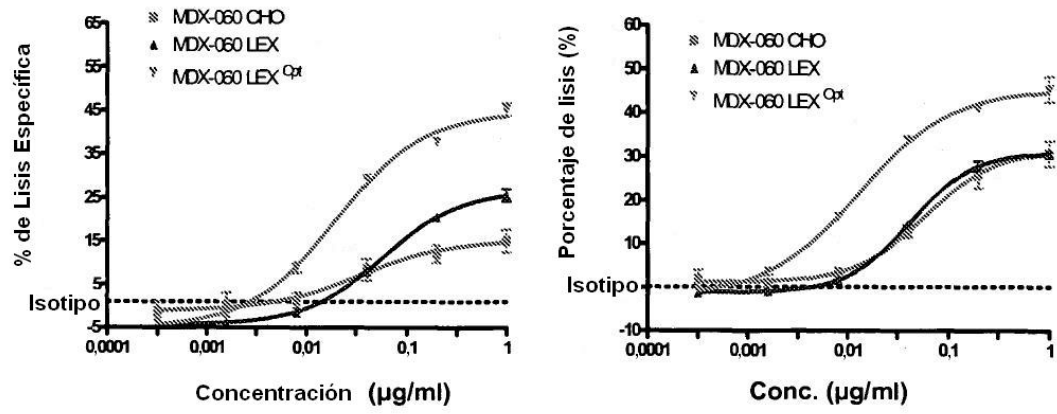
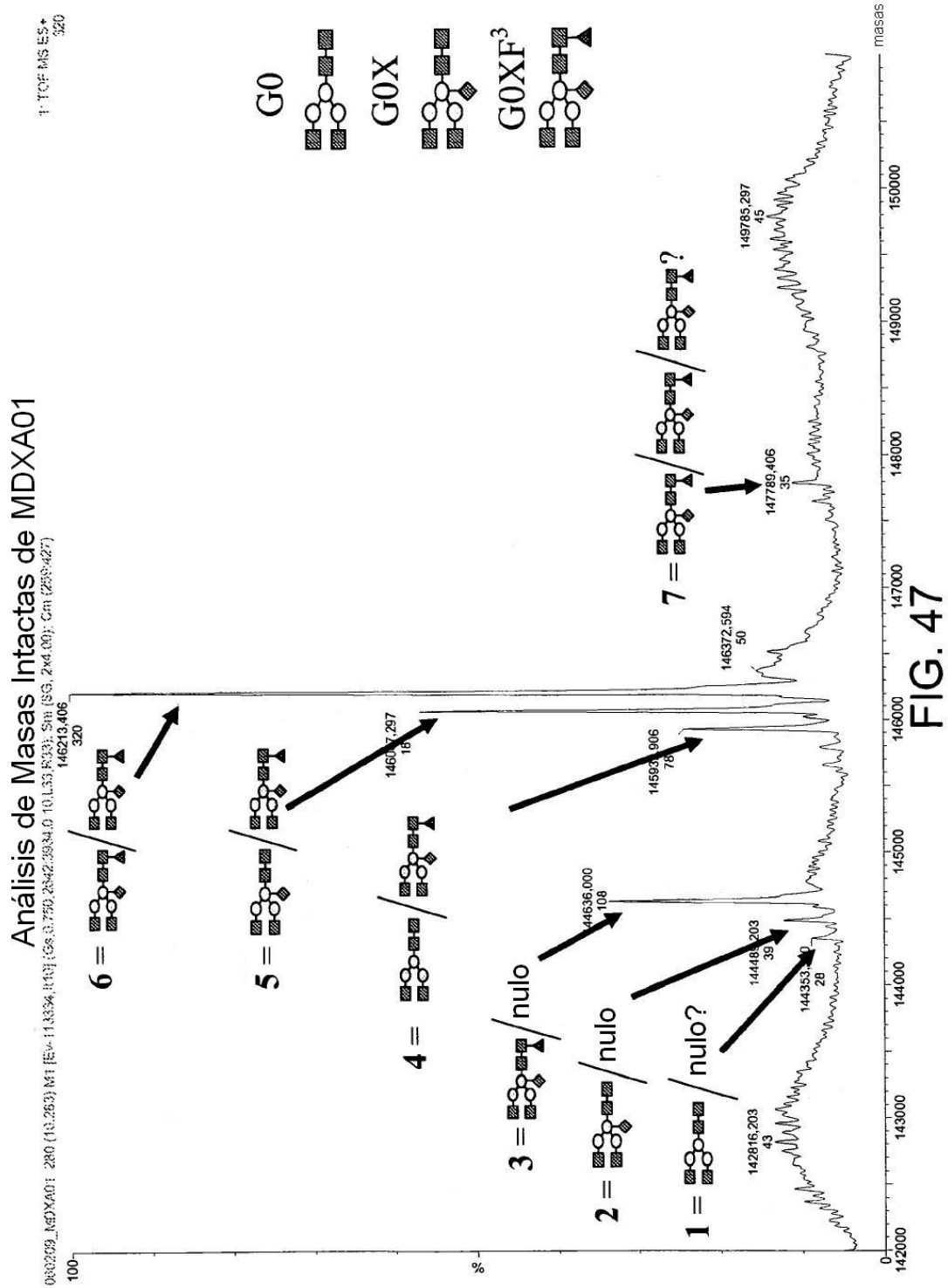


FIG. 46



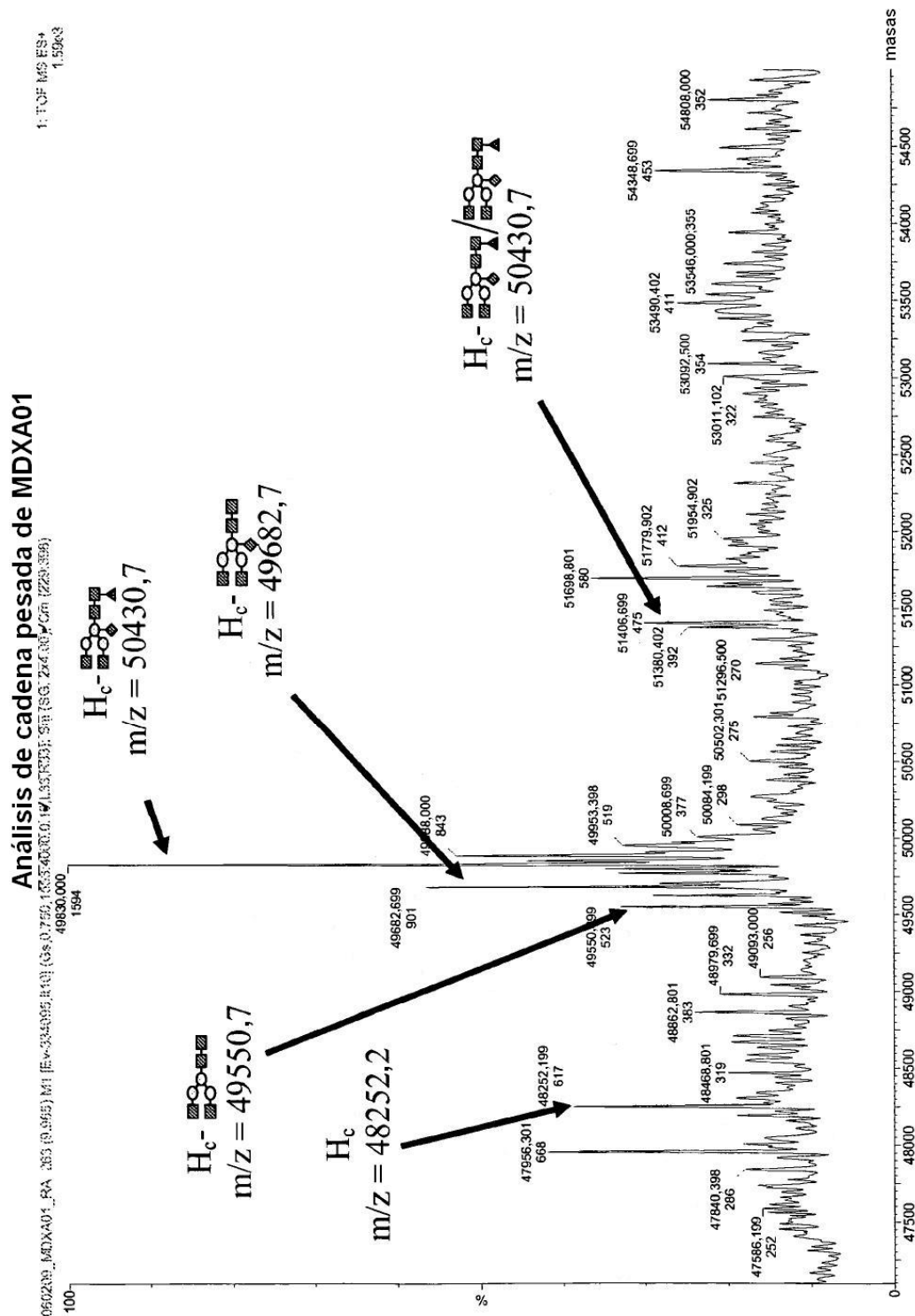


FIG. 48

Análisis de Masas Intactas de MDXA04

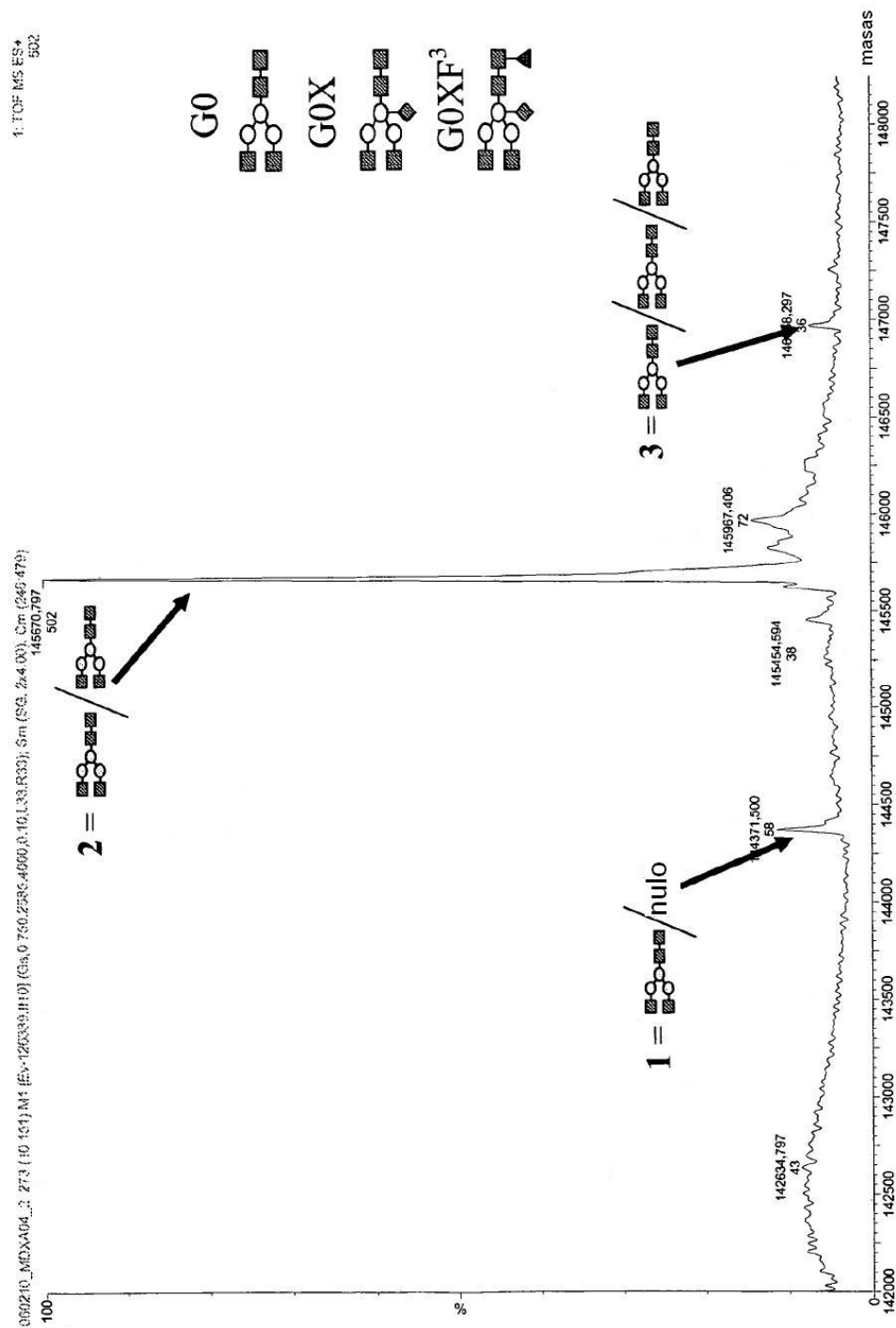


FIG. 49

Análisis de cadena pesada de MDXA04

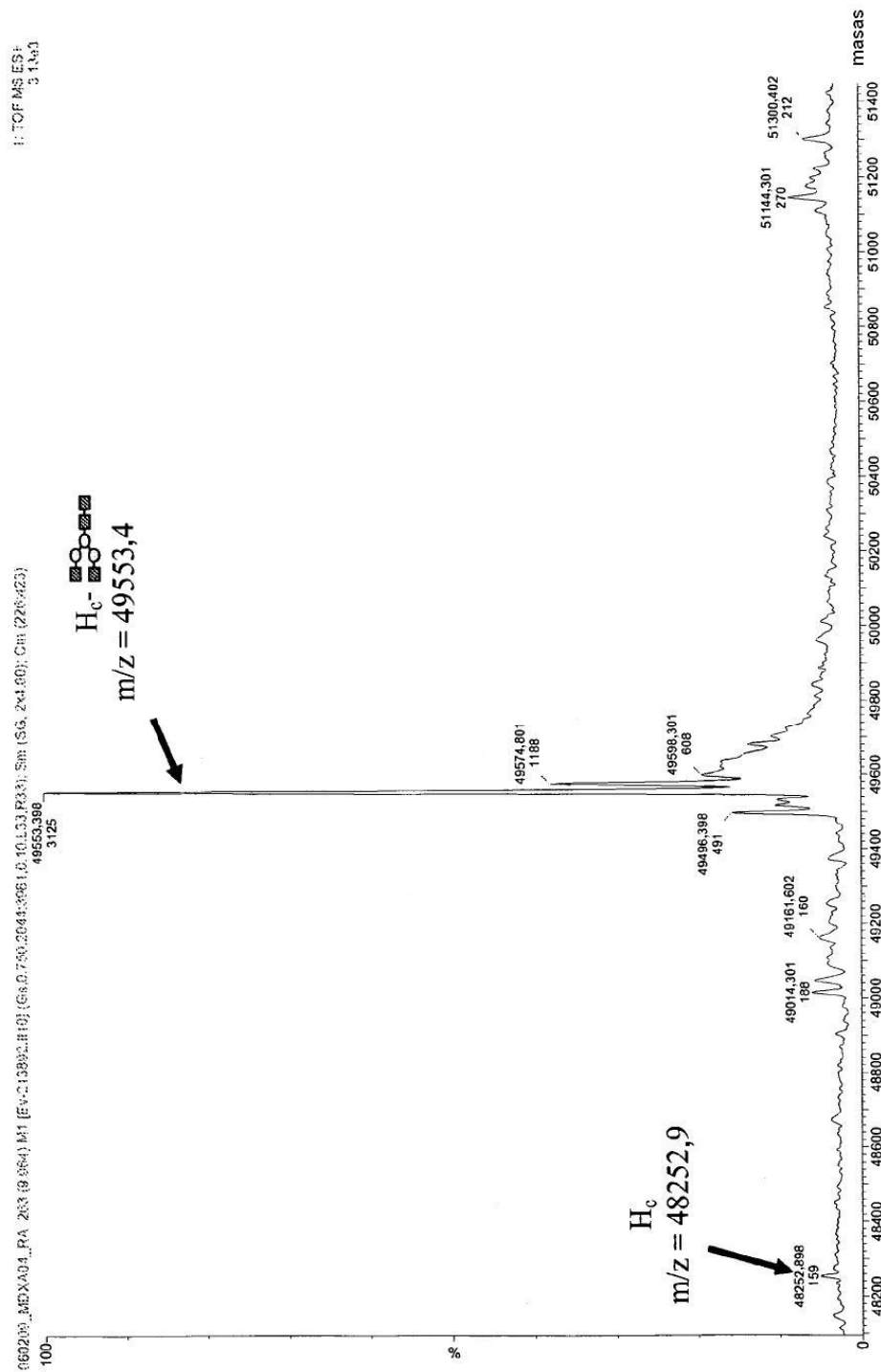


FIG. 50

Análisis MALDI de N-glicanos de la línea mAb104 escalada

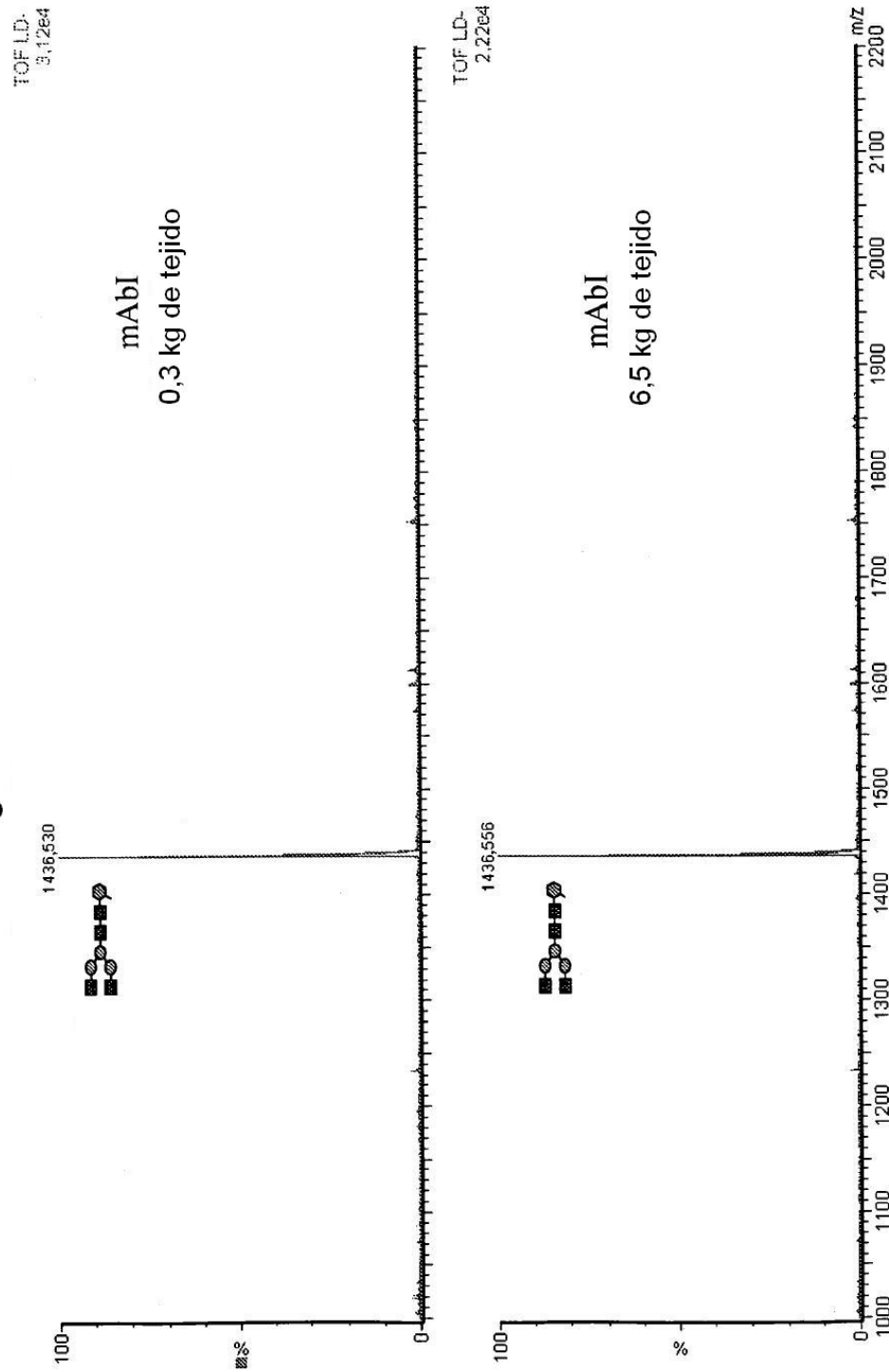


FIG. 51A

Análisis de fluorescencia de HPLC de N-glicanos de línea mAb104 escalada

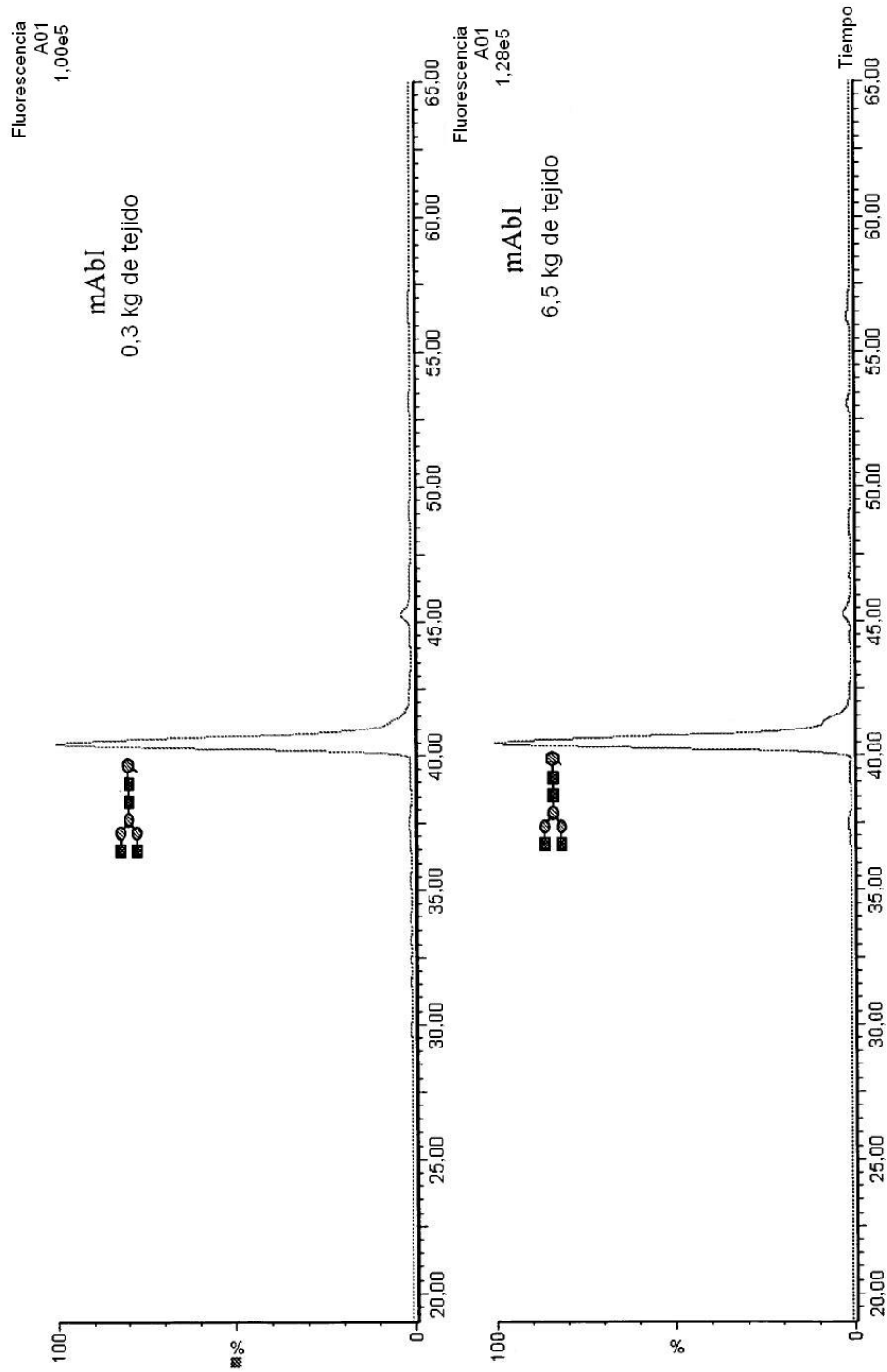


FIG. 51B

Análisis MALDI de N-glicanos de glicoproteína endógena de *Lemna* en líneas naturales y de ARNi

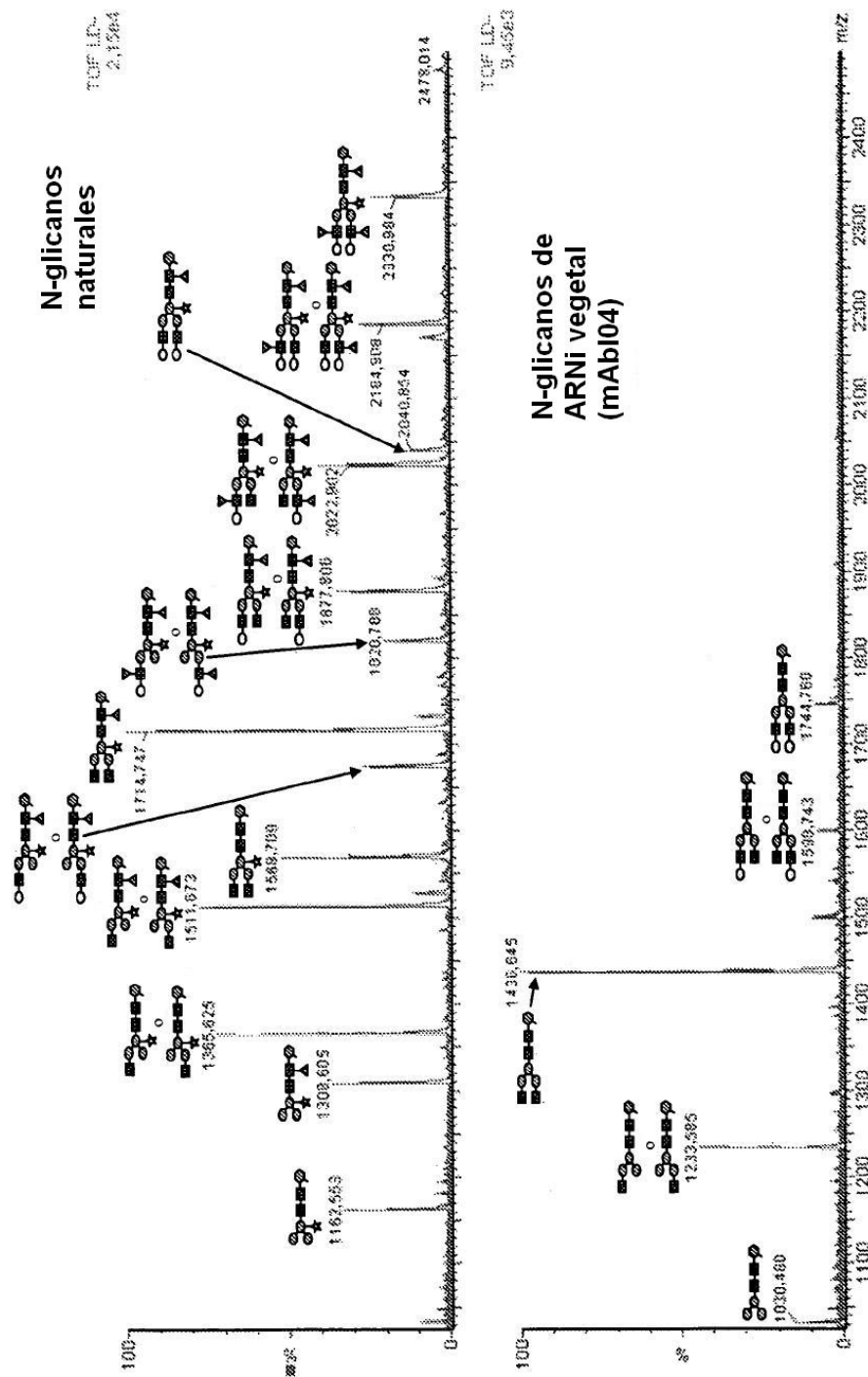


FIG. 52

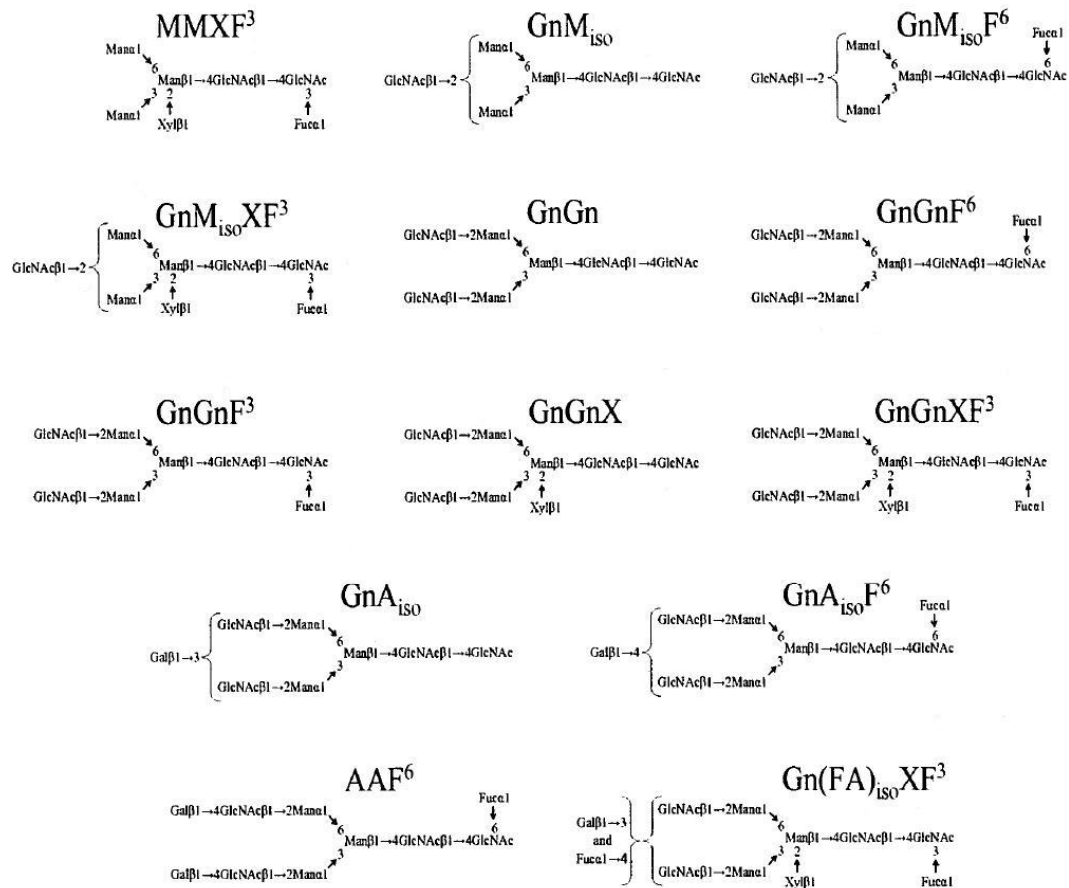


FIG. 53